

Univerzita Palackého v Olomouci

Přírodovědecká fakulta

Katedra geoinformatiky

**ANALÝZA FAKTORŮ OVLIVŇUJÍCÍCH
VZNIK CHYB V DEM**

Magisterská práce

Bc. Veronika OBADÁLKOVÁ

Vedoucí práce: RNDr. Jana Svobodová Ph.D.

Olomouc 2015

Geoinformatika

ANOTACE

Diplomová práce se zabývá nalezením hlavních parametrů, které lze odvodit z DEM a z povahy vstupních dat (např. rozmístění či hustota dat) a lze je použít pro odhad chyb v DEM. Dalším, neméně důležitým výsledkem, je sestavení regresních modelů umožňujících odhad chyby v DEM na základě těchto vybraných parametrů. Pro tento účel jsou výsledky založené na uměle generovaných datech a na reálných datech z leteckého laserového skenování části obce Vysoké Pole.

K hodnocení kvality DEM byly využity jak neprostorové (globální), tak prostorové (lokální) metody hodnocení. Neprostorové hodnocení kvality DEM bylo založeno na výpočtu váženého pořadí dle vybraných neprostorových ukazatelů (střední kvadratické chyby, celkové absolutní chyby, minimální a maximální chyby, průměru a směrodatné odchylky). K prostorovému hodnocení kvality DEM byla zvolena metoda lokální shlukové analýzy LISA. Součástí práce byl i výzkum vztahů mezi hodnotami chyb a hodnotami zkoumaných parametrů pomocí korelační a regresní analýzy.

Výsledky práce by se měly stát návodem, jak vytvořit kvalitní DEM z hlediska metrické přesnosti a také jak nalézt místa (shluky) se statisticky významnými hodnotami vysokých chyb.

KLÍČOVÁ SLOVA

digitální výškový model (DEM); chyby v DEM; kvalita DEM; korelační analýza; regresní analýza

Počet stran práce: 62

Počet příloh: 8

ANOTATION

This diploma thesis deals with the finding of the main parameters that can be derived from DEM characteristics and assessment quality of DEM from input data (e.g. position and density of input data) and can be used to estimate the errors in the DEM. Another, equally important result, is building of regression models allowing estimation errors in the DEM on the basis of the selected parameters. For this purpose, the results are based on artificially generated data and on real data from laser scanning of the part of municipality Vysoké Pole.

The non-spatial (global) and spatial (local) assessment methods were employed for the assessment quality of DEM. The non-spatial assessment of the DEM quality was based on the calculation of the weighted order according to the chosen non-spatial indicators (root mean square error, total absolute error, minimum and maximum error, mean and standard deviation). The method of local cluster analysis LISA has been chosen for the spatial assessment of the DEM quality. Another part of the thesis consists in the research of the relationships between the error values and the values of researched parameters. These relationships were examined by means of correlation and regressive analysis.

The results of this thesis should help to create a high quality DEM from the metric accuracy point of view, to find places (clusters) with statistically significant values of high error values.

KEYWORDS

digital elevation model (DEM); errors in DEM; quality of DEM; correlation analysis; regression analysis

Number of pages: 62

Number of appendixes: 8

Prohlašuji, že

- diplomovou práci, včetně příloh, jsem vypracovala samostatně a uvedla jsem všechny použité podklady a literaturu,

- jsem si vědoma, že na moji diplomovou práci se plně vztahuje zákon č.121/2000 Sb. - autorský zákon, zejména § 35 – využití díla v rámci občanských a náboženských obřadů, v rámci školních představení a využití díla školního a § 60 – školní dílo,

- beru na vědomí, že Univerzita Palackého v Olomouci (dále UP Olomouc) má právo nevýdělečně, ke své vnitřní potřebě, diplomovou práci užívat (§ 35 odst. 3),

- souhlasím, aby jeden výtisk diplomové práce byl uložen v Knihovně UP k prezenčnímu nahlédnutí,

- souhlasím, že údaje o mé diplomové práci budou zveřejněny ve Studijním informačním systému UP,

- v případě zájmu UP Olomouc uzavřu licenční smlouvu s oprávněním užít výsledky a výstupy mé diplomové práce v rozsahu § 12 odst. 4 autorského zákona,

- použít výsledky a výstupy mé diplomové práce nebo poskytnout licenci k jejímu využití mohu jen se souhlasem UP Olomouc, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly UP Olomouc na vytvoření díla vynaloženy (až do jejich skutečné výše).

V Olomouci dne

Veronika Obadálková

Na tomto místě bych ráda poděkovala RNDr. Janě Svobodové, Ph.D. za odborné rady, cenné připomínky, ochotu a čas, který mi byla ochotna věnovat při tvorbě diplomové práce. Také bych chtěla poděkovat mé rodině a přátelům za podporu během studia.

Vložený originál **zadání** magisterské práce (s podpisy vedoucího katedry, vedoucího práce a razítkem katedry). Ve druhém výtisku práce je vevázána fotokopie zadání.

OBSAH

SEZNAM POUŽITÝCH ZKRATEK	9
ÚVOD	10
1 CÍLE PRÁCE.....	11
2 POUŽITÉ METODY A POSTUPY ZPRACOVÁNÍ	12
2.1 Použitá data	12
2.2 Použité programy	12
2.3 Postup zpracování	13
3 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY	16
3.1 Digitální výškový model	16
3.1.1 Chyby v DEM.....	18
3.2 Faktory ovlivňující vznik chyb v DEM	19
3.2.1 Metody sběru vstupních dat.....	19
3.2.2 Metody vzorkování	20
3.2.3 Typy datového modelu	21
3.2.4 Metody interpolace	22
3.2.5 Morfometrické charakteristiky	23
3.3 Metody hodnocení přesnosti	25
3.4 Hodnocení vztahů mezi parametry.....	28
4 TVORBA DEM A ODVOZENÍ ZKOUMANÝCH PARAMETRŮ.....	31
4.1 Získání a úprava vstupních dat.....	31
4.2 Tvorba DEM	33
4.3 Základní analýza dat	33
4.4 Odvození zkoumaných parametrů.....	35
5 HODNOCENÍ KVALITY DEM ZE VSTUPNÍCH DAT.....	36
6 ANALÝZA VZTAHŮ MEZI HODNOTAMI CHYB A HODNOTAMI ZKOUMANÝCH PARAMETRŮ	41
6.1 Korelační analýza matematického modelu	41
6.2 Regresní analýza matematického modelu	45
6.2.1 Vícenásobná lineární OLS regrese – metoda nejmenších čtverců.....	45
6.2.2 Geograficky vážená regrese - GWR.....	51
7 VERIFIKACE VÝSLEDKŮ NAD DATY Z LETECKÉHO LASEROVÉHO SKENOVÁNÍ.....	53
7.1 Korelační analýza.....	53
7.2 Regresní analýza	53
7.2.1 Vícenásobná lineární OLS regrese – metoda nejmenších čtverců.....	53
7.2.2 Geograficky vážená regrese - GWR.....	57

8	VÝSLEDKY	58
9	DISKUZE	60
10	ZÁVĚR	61
	POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE	
	PŘÍLOHY	

SEZNAM POUŽITÝCH ZKRATEK

Zkratka	Význam
AE	Absolute Error
AIC	Akaike Information Criterion
DEM	Digital Elevation Model
GIS	Geographic Information System
GWR	Geographically Weighted Regression
LAS	Lidar Archive Standard
LIDAR	Light Detection and Ranging
LISA	Local Indicators of Spatial Association
OLS	Ordinary Least Square
RMSE	Root Mean Square Error
RSS	Residual Sum of Squares
SD	Standard Deviation
SHP	Shapefile
VIF	Variance Inflation Factor

ÚVOD

Reliéf terénu naší planety hraje významnou úlohu v mnoha vědních disciplínách týkajících se nejen nauky o Zemi. Z vědních oborů to jsou například geomorfologie, hydrologie, geologie, ekologie, geografie aj. S příchodem výpočetní techniky a umožnění tvorby digitálních modelů reliéfu v digitálním prostředí počítačů dostala tahle problematika ještě větší význam. Geografické informační systémy (GIS) nejsou jen softwarem pro odborníky, ale především důležitým nástrojem pro širokou veřejnost. Slouží pro podporu rozhodování pracovníků územního plánování, vodohospodářství, stavebnictví apod. Dle Voženilka a kol. (2001) jsou jejím základem prostorové analýzy, pomocí nichž lze ze vstupních dat odvodit nová data popisující jevy a procesy na zemském povrchu a získat tak nové informace. Kvalita těchto výsledných dat i odvozených informací však závisí z velké části na kvalitě vstupních dat.

Digitální výškový model je zjednodušením zemského povrchu v prostředí GIS. Kvalita DEM je primárně ovlivněna mnoha faktory, jakými jsou metody sběru vstupních dat, metody vzorkování, typy datového modelu a metody interpolace. Jedním z nejdůležitějších faktorů lze však označit vliv rozmístění vstupních bodů, jejichž výběr se provádí různými metodami vzorkování.

Tato práce pojednává o přesnosti digitálních výškových modelů různých typů reliéfu. Předmětem práce je tedy popsat postup hodnocení kvality digitálních výškových modelů z hlediska povahy vstupních dat (např. rozmístění či hustoty dat) a také vliv kvality digitálních výškových modelů na hodnoty zkoumaných parametrů, kterými jsou sklony svahu, orientace svahu, křivosti a jiné.

Tyto poznatky byly využity při plánování a tvorbě diplomové práce.

1 CÍLE PRÁCE

Cílem diplomové práce je nalézt všechny důležité parametry, které lze odvodit z DEM a z povahy vstupních dat (např. rozmístění či hustota dat) a lze je použít pro odhad chyb v DEM.

Po nastudování potřebné literatury, zabývající se problematikou hodnocení kvality digitálních výškových modelů, budou v teoretické části zpracovány důležité parametry, které lze použít pro odhad chyb v DEM. Druhá část práce bude praktická a bude založena na opakovaném vytváření regresních modelů s různými vysvětlujícími parametry (nutná analýza reziduí). Po nalezení všech důležitých parametrů bude sestaven regresní model umožňující odhad chyby v DEM.

Samotné testování bude prováděno nad vstupními daty, kterými budou uměle generovaná data simulující různé typy dat využívané pro tvorbu DEM a reálná data z leteckého laserového skenování části obce Vysoké Pole. Výsledky získané testováním budou dále vyhodnoceny pomocí statistických metod a opatřeny komentářem. Na základě výsledků pak budou vysloveny závěry o analyzovaných parametrech a stanovení jejich významnosti pro modelování chyb v DEM.

2 POUŽITÉ METODY A POSTUPY ZPRACOVÁNÍ

Na začátku diplomové práce bylo nutné podrobně nastudovat problematiku hodnocení kvality digitálních výškových modelů. Následně byly vybrány vhodné parametry, které lze použít pro odhad chyby v DEM a ideální postupy zpracování práce.

2.1 Použitá data

Pro účely diplomové práce byla použita, jako testovací datová sada, umělá data náhodně generovaná v programu R, kde bylo vygenerováno 40 000 bodů v pravidelné mřížce. Každý bod obsahoval souřadnici x a y a výšku z . Dále bylo třeba ověřit funkčnost regresního modelu na reálných datech, a proto byla zvolena data z leteckého laserového skenování obsahující část katastrálního území obce Vysoké Pole na Zlínsku ve formátu *.las. Laserová data byla zakoupena od společnosti GEODIS Brno, spol. s. r.o. Katedrou geoinformatiky Univerzity Palackého v Olomouci. Pořízena byla dne 27. 6. 2012 skenerem Leica ALS50-II při průměrné výšce letu 1 000 m a při maximální rychlosti letounu 241 km/h. Data se skládají z pěti letových řad s nominálním příčným překrytem 35% a úhlem skenování 25°. Reálná hustota všech pořízených bodů je 5,272 bodů na metr čtvereční, z toho je 3,229 bodů terénu. Výsledná laserová data jsou klasifikována do tříd dle standardu ASPRS (American Society for Photogrammetry and Remote Sensing) (viz tab. 1).

Tab. 1 Klasifikace bodů ve sledované oblasti dle ASPRS standardu

(zdroj: http://www.asprs.org/a/society/committees/standards/asprs_las_format_v12.pdf)

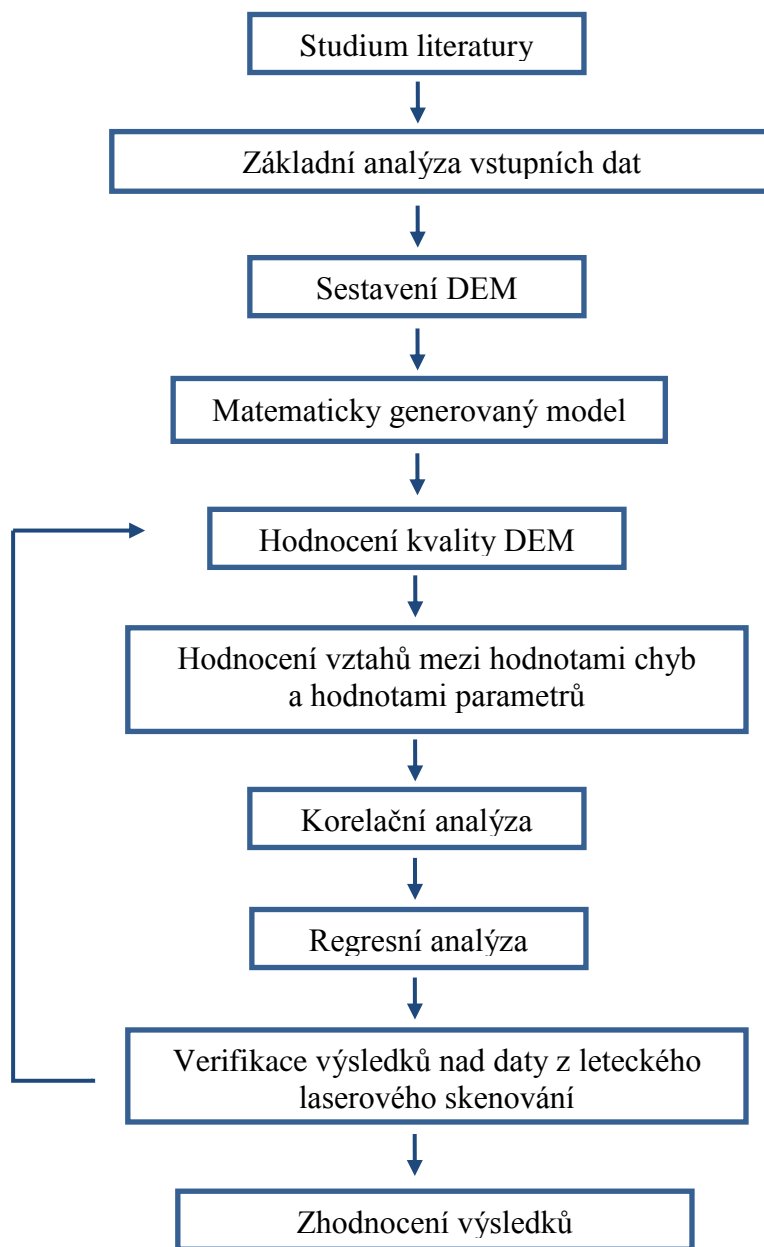
Číslo kategorie	Význam
1	neklasifikováno
2	zemský povrch
3	nízká vegetace
4	střední vegetace
5	vysoká vegetace
6	budovy
7	nízké body (šum)

2.2 Použité programy

Ke zpracování uměle generovaných dat a jejich analýz a dat z leteckého laserového skenování bylo využito programu ArcGIS 10.1 od společnosti ESRI. K dílčímu statistickému zpracování dat byl využit software R v. 2.15.3. Pro grafickou vizualizaci a popis všech výsledků bylo využito kancelářského balíku Microsoft Office 2010 od společnosti Microsoft.

2.3 Postup zpracování

V následující kapitole je uveden postup zpracování diplomové práce (viz obr. 1) a přehled použitých metod.



Obr. 1 Postup zpracování diplomové práce.

V první fázi bylo třeba zpracovat umělá data náhodně generovaná a reálná data leteckého laserového skenování a vytvořit tak přesný digitální výškový model. Pro tvorbu DEM byla použita interpolační metoda Topo to Raster, která kombinuje několik interpolačních metod, aby vytvořila hydrologicky korektní model terénu a je implementovaná v prostředí ArcGIS Desktop. Pro každou datovou sadu (viz tab. 2) bylo vygenerováno 9 gridů (celkem 18 gridů) s rozlišením pixlu 5 m a s různou hustotou

a rozmístěním vstupních bodů – náhodně rozmístěné body o různých hustotách, pravidelně rozmístěné body a vrstevnice (viz tab. 3). Jako referenční grid, který sloužil pro výpočet některých chyb, byl zvolen originální 100% grid. Po vytvoření gridů s různým rozmístěním a hustotou vstupních bodů z nich bylo třeba vypočítat hodnoty zkoumaných parametrů a to nadmořské výšky, převýšení, vzdálenosti ke vstupním bodům (u vrstevnic k vertexům), sklonu svahu, orientace svahu, orientace reklasifikované, celkové, planární a profilové křivosti, dále extremity minima, maxima a průměru a také parametry získané pronásobením základních morfometrických parametrů sklonu svahu × profilové křivosti a orientace svahu × planární křivosti. Celkem vzniklo 252 gridů. Všechny výpočty proběhly v prostředí ArcGIS Desktop.

Tab. 2 Datová sada použitá pro hodnocení kvality DEM

	Velikost území [km]	Rozlišení pixlu [m]	Rozsah hodnot [m n. m.]	
			min	max
Matematicky generovaný model	1 x 1	5	-19	16
Laser scan oblasti Vysoké Pole	1,1 x 1,4	5	398	506

Tab. 3 Přehled rozmístění a hustoty vstupních bodů

Rozmístění vstupních bodů	Hustota vstupních bodů
vrstevnice	1 m, 3 m, 5 m
náhodné	5 %, 10 %, 55 %
pravidelné	15 m, 25 m, 50 m

Podstatnou částí práce bylo hodnocení kvality výškových modelů jak pomocí globálních neprostorových, tak lokálních prostorových charakteristik. Pro území ke srovnání DEM, podle neprostorových charakteristik, bylo využito *vážené pořadí* založené na výpočtu *střední kvadratické chyby (RMSE)*, *celkové absolutní chyby (AE)*, *minimální a maximální chyby*, *průměru a směrodatné odchylky (SD)* dle vzorců uvedených v kapitole 3.3. Pro každou z uvedených charakteristik bylo vytvořeno vážené pořadí a díky němu lze DEM seřadit od nejkvalitnějšího po méně kvalitní. Pro prostorové hodnocení DEM bylo využito *lokální shlukové analýzy LISA*, kde výsledkem je vymezení shluků statisticky významných chyb a pomocí níž lze sledovat prostorové rozmístění chyb.

Další podstatnou částí práce bylo zjistit, zda existuje závislost mezi velikostí chyb v DEM a hodnotami zkoumaných parametrů. K tomu slouží korelační a regresní analýza.

Nejprve bylo nutné provést korelační analýzu, kde bylo zjištěno, zda mezi proměnnými (nadmořská výška, převýšení, vzdálenost ke vstupním bodům (u vrstevnic k vertexům), sklon svahu, orientace svahu, orientace svahu reklasifikovaná, planární, profilová a celková křivost, extremita minima, maxima a průměru a také parametry získané pronásobením základních morfometrických parametrů sklonu svahu \times profilové křivosti a orientace svahu \times planární křivosti) existuje či neexistuje závislost (korelace). Byl použit jednoduchý typ korelace založený na výpočtu *Spearmanova pořadového korelačního koeficientu*. V případě, že závislost existuje, je možné ji vyjádřit pomocí regresní analýzy. Pro porovnání byla využita *vícenásobná lineární OLS regrese*, která byla využita k vyjádření vztahů mezi hodnotami chyb jako závisle proměnnou a hodnotami zkoumaných parametrů jako nezávisle proměnnými, tak i *geograficky vážená regrese (GWR)*.

Poslední fází této magisterské práce byl po nalezení všech důležitých parametrů sestaven regresní model umožňující odhad chyby v DEM.

3 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY

Cílem této kapitoly je uvést problematiku týkající se digitálních výškových modelů. V současné době není mnoho autorů, kteří se zabývají výzkumem vztahu mezi hodnotami chyb a hodnotami morfometrických parametrů pomocí korelační a regresní analýzy. Jako hlavní autory, zabývající se tímto výzkumem, můžeme zmínit Carlisle (2000, 2002, 2005), Erdogana (2009, 2010) a Evanse (1972).

Erdogan (2009) ve své práci použil 4 základní morfometrické parametry pro regresní analýzu a těmi byly: sklon, orientace, celková křivost a nadmořská výška. Evans (1972) zase nahradil celkovou křivost za planární a profilovou a spolu s dalšími parametry je považuje za dostačující pro zcela vyčerpávající množství povrchových tvarů. Carlisle (2002, 2005) se ve své práci neomezuje jen na těchto 6 základních morfometrických parametrů, ale testuje i použití dalších proměnných a to relativní výškovou členitost, texturu, extremitu průměru, extremitu minima, extremitu maxima a vzdálenost k vertexům. Pro většinu parametrů testoval i druhé a třetí mocniny jejich proměnných, s cílem zjistit nelineární vztahy mezi hodnotami chyb a hodnotami morfometrických parametrů.

Pro posouzení základních vztahů mezi hodnotami chyb a hodnotami zkoumaných parametrů byl použit výpočet korelačního koeficientu zvláště pro každý parametr. Výše uvedení autoři (např. Carlisle 2005, Erdogan 2010) uvádějí, že není vhodné zkoumat sílu vztahů jednotlivých parametrů a chyb samostatně. Proto je třeba použít více nezávislých (vysvětlujících) parametrů zároveň a toho lze dosáhnout za použití regresní metody.

3.1 Digitální výškový model

Reliéf terénu naší planety hraje významnou úlohu v mnoha vědních disciplínách týkajících se nejen nauky o Zemi. Z vědních oborů to jsou například geomorfologie, hydrologie, geologie, ekologie, geografie aj. S příchodem výpočetní techniky a umožnění tvorby digitálních modelů reliéfu v digitálním prostředí počítačů dostala tato problematika ještě větší význam. Podle Krcha (1990) je **digitální model reliéfu (DMR)** – angl. *Digital Model of Relief* - „reprezentativní soubor bodů reliéfu terénu vybraných podle určitých pravidel, polohově lokalizovaných a s přiřazeným vektorem (sloupcem hodnot) parametrů reliéfu terénu. Jde tedy o body, informace o nich a pravidla používání těchto informací“.

Prostorové modely zemského povrchu můžeme rozdělit následovně (Bayer, 2008):

- Digitální model terénu (DTM), někdy také digitální model reliéfu (DMR)
- Digitální výškový model (DEM)
- Digitální model povrchu (DSM)

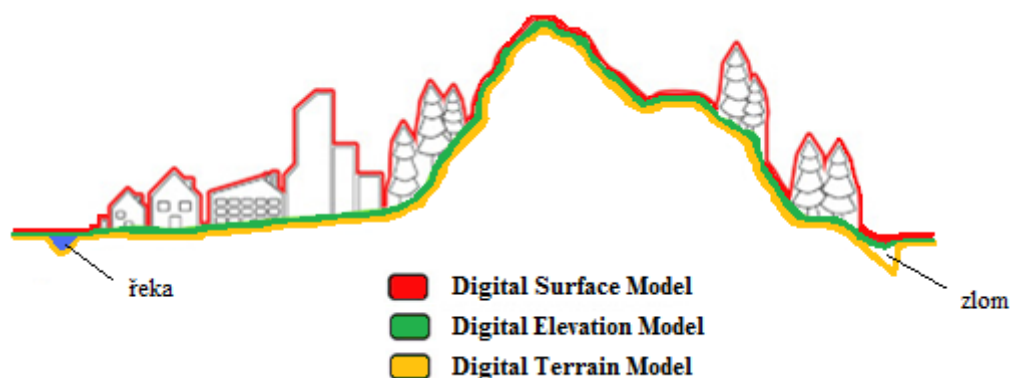
Mezi pojmy v anglickém a českém znění však panuje značná nejednotnost, proto je tedy vhodné na začátek vymezit definice jednotlivých pojmů a zkratk.

Termíny DTM a DEM jsou si podobné a v praxi často zaměňovány. Oba modelují čistě holý terén zemského povrchu bez vegetace a bez lidských výtvorů jako jsou budovy, mosty, silnice atd. (Voženílek a kol., 2001). Ale i mezi nimi existují malé rozdíly.

Digitální model terénu (DTM) – angl. *Digital Terrain Model* – má širší význam než DEM, do jeho tvorby kromě výškových dat vstupují i geografické prvky a přírodní jevy jako jsou řeky a zlomové linie. Může také obsahovat odvozené údaje o terénu jako je sklon svahu, orientace, zakřivení, viditelnost, atd. (viz obr. 2) (Oksanen, 2006; Li et al., 2005).

Digitální výškový model (DEM) – angl. *Digital Elevation Model* – je oproti tomu vytvořen pouze z výškových dat (z výškových bodů nebo vrstevnic) o známých nadmořských výškách vztažených k rovinným nebo zeměpisným souřadnicím (Oksanen, 2006; Li et al., 2005).

Digitální model povrchu (DSM) – angl. *Digital Surface Model* – podle Vojtka (2008) zachycuje průběh topografické plochy georeliéfu, který je doplněn o prvky tvořící jak jeho přírodní, tak jeho antropogenní pokryv (například stromy, budovy, mosty apod.).



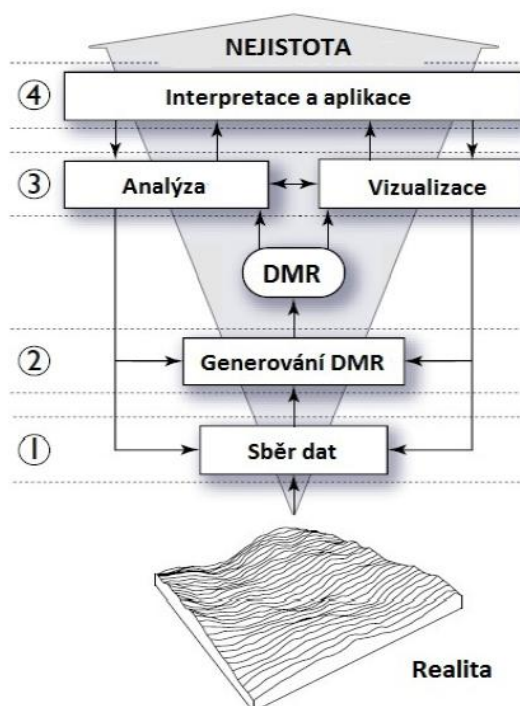
Obr. 2 Jednotlivé digitální modely zemského povrchu

(zdroj: upraveno dle <http://www.computamaps.com/newsletter/3-3/newsletter3-3.html>).

Stejně jako ostatní soubory prostorových dat, tak i data digitálních výškových modelů podléhají chybám. Chyby (angl. errors) jsou běžně definovány jako rozdíl mezi skutečnými a modelovanými hodnotami.

V oblasti tvorby digitálního výškového modelu (DEM) lze snadno demonstrovat příkladný vznik chyb. Což ve výsledku dává vzniknout nejistotě (angl. uncertainty), jejíž míra může postupně gradovat už v průběhu celého pracovního procesu (viz obr. 3). Na počátku veškerých prací je samozřejmě sběr dat a již zde se vyskytuje prvotní zárodek nejistoty, zapříčiněné přístrojem a lidským faktorem, který bývá dále jen umocňován. Přispívají k tomu postupně jak metody pro generování a tvorbu modelu, tak i další použité analýzy a vizualizace. Na konci stojí logicky interpretace a aplikace získaného

výsledku, jenže i zde se může vyskytnout řada chyb zapříčiněna komunikační propastí mezi uživatelem a počítačem, které mohou podtrhnout výslednou nejistotu celého procesu tvorby (Oksanen, 2006).



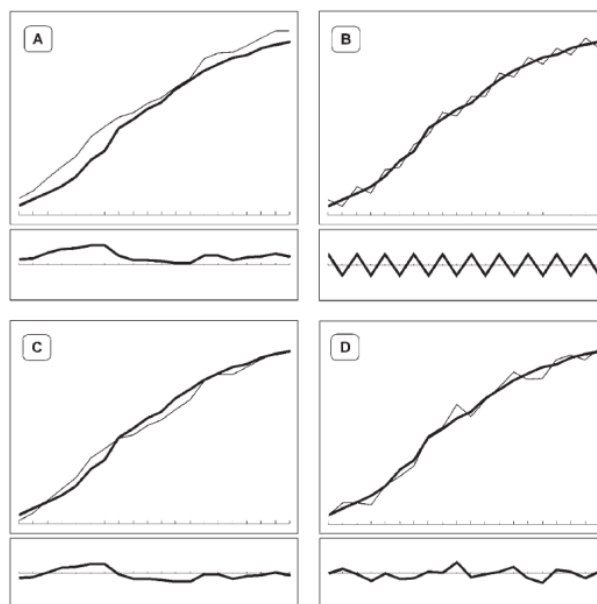
Obr. 3 Hlavní části modelování DEM a fáze, při kterých může nejistota vstupovat do procesu tvorby DEM (zdroj: přeloženo dle Oksanen, 2006).

3.1.1 Chyby v DEM

Chyby v DEM se mohou vyskytovat v horizontálním i vertikálním směru. Nicméně v diplomové práci se více zaměříme na vertikální, tedy výškové záznamy. Tyto chyby, které jsou popsány níže, mohou negativně ovlivnit další výpočet zkoumaných parametrů. Podle Fischera a Tatea (2006) lze původ chyb zařadit do tří charakteristických kategorií (viz obr. 4):

- **Hrubé chyby** (angl. gross errors/blunders), které mohou mít svůj původ například v selhání lidského faktoru či použité aparatury. Avšak hrubé chyby lze poměrně snadno rozpoznat i odstranit
- **Systematické chyby** (angl. systematic errors), které se vyskytují podle určitého pravidla (viz obr. 4 - A, B). Jejich příkladem mohou být tzv. duchové, neboli obrysy linií identifikovatelné v mnoha DEM odvozených od vrstevnicových dat, nebo také známé umělé terasy. Systematické chyby mohou být eliminovány nebo podstatně redukovány
- **Náhodné chyby** (angl. random errors) jsou čistě náhodné a nepředvídatelné, pochází z vysoké rozmanitosti měření, nelze je zcela vhodně modelovat a bývají

konceptně reprezentovány náhodnými variacemi okolo střední hodnoty (viz obr. 4 - C, D)



Obr. 4 Porovnání profilu DEM a výskytu chyb. Výskyt: (A, B) systematických chyb, (C) prostorově autokorelovaných chyb (normální stav), (D) náhodných chyb (bez prostorové autokorelace). V každé situaci znázorňuje horní obrázek povrch Země silnou čarou a povrch s chybami čarou tenkou, spodní obrázek zobrazuje samotnou chybu (zdroj: Fisher a Tate, 2006).

3.2 Faktory ovlivňující vznik chyb v DEM

Cílem této kapitoly je uvést základní faktory, které primárně ovlivňují kvalitu DEM. Jsou za ně považovány metody sběru dat, metody určující výběr rozmístění vstupních dat (metody vzorkování), výběr datového modelu a interpolační metody.

3.2.1 Metody sběru vstupních dat

Metody sběru vstupních dat ovlivňují kvalitu digitálních výškových modelů z hlediska metrické přesnosti, proto byly v diplomové práci zvoleny dvě testovací datové sady a to umělá data náhodně generovaná a reálná data leteckého laserového skenování, kde se předpokládá vysoká přesnost.

Pro pořizování dat je k dispozici několik různých metod. V závislosti na jejím výběru lze již dopředu odhadnout typy chyb, které se v DEM v důsledku použité metody mohou vyskytovat. Při volbě metody vhodné pro zpracování vybrané oblasti musíme zohlednit kritéria jednotlivých metod: velikost území, požadovaná polohová a výšková přesnost, požadovaná maximální hustota bodů, čas a cena. Volba nejvhodnější metody závisí především na konkrétním postupu a účelu zpracování a lze ji rozdělit do tří skupin (Svobodová, 2011):

- **Z dálkového průzkumu Země** (fotogrammetrie, radarové snímání, laserové snímání), kdy výhodou je pořizování dat na dálku i pro rozsáhlá území s menší časovou a i finanční náročností. U těchto metod je důležitým krokem zpracování surových dat před tím, než budou použita pro tvorbu DEM
- **Z pozemního měření** (geodetická měření, globální polohové systémy (GPS)), kdy nevýhodou je časová a finanční náročnost jejich pořízení, ale za to produkují velmi přesná data použitelná přímo pro tvorbu DEM
- **Z digitalizace existujících analogových dat**, kdy přesnost závisí na metodách sběru původních dat

3.2.2 Metody vzorkování

S pořizováním geodat pro digitální modely úzce souvisí problematika vzorkování. Povrch se skládá z nekonečného množství bodů, takže cílem metody vzorkování je tedy popsat spojitý jev, jakým je povrch terénu, pomocí konečného počtu vzorků (bodů) a určit jejich optimální rozmístění. Na kvalitu digitálního modelu mají největší vliv hustota a rozložení bodů, prahová hodnota odchylky výšky reliéfu od výšky interpolované z digitálního modelu a aproximace reliéfu strukturními a lomovými liniemi. U této metody může dojít k převzorkování, kdy je povrch reprezentován vzorky dokonale, ale za cenu vysokého množství dat a k podvzorkování, kdy by mohlo dojít k výraznému zkreslení povrchu (Rapant, 2006).

Metodami vzorkování je dána hustota vstupních dat (včetně jejich rozmístění) v návaznosti na členitost reliéfu. Jejich hustota působí na shlazení výsledného DEM. Malá hustota vstupních dat má za následek větší míru vyhlazení výstupního DEM, které může, ale nemusí být vždy žádoucí. Naopak příliš velká hustota vstupních dat s sebou nese riziko zbytečné redundance dat a vyšší výpočtové náročnosti (Svobodová, 2011).

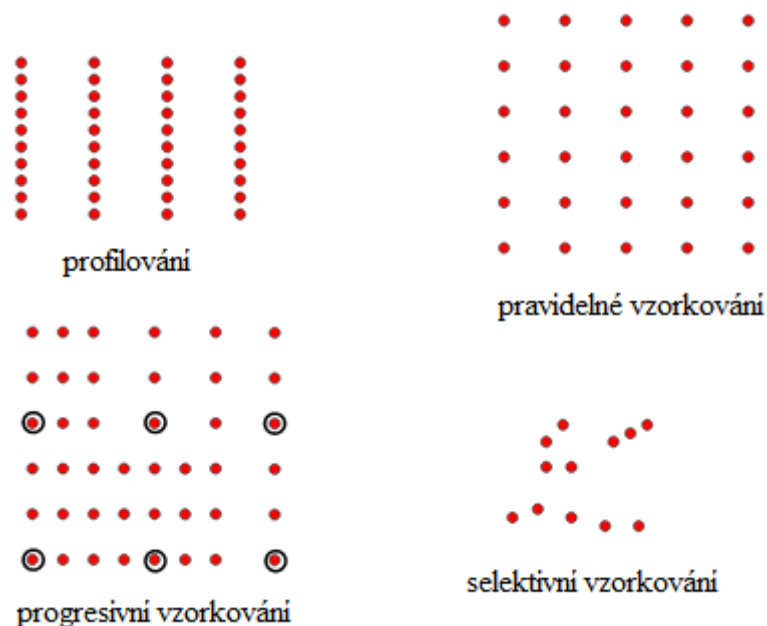
Pro určení optimální hustoty bodů podle základní vzorkovací teorie lze velikost kroku měření odvodit ze vztahu:

$$\Delta l \leq 0,5l_{min} \quad (1)$$

kde Δl je krok měření a l_{min} je minimální velikost objektu reálného světa (respektive nejmenšího detailu), který chceme ve výsledném vzorku zachytit (Rapant, 2006).

Nejčastěji používané postupy při vzorkování (obr. 5):

- pravidelné vzorkování,
- nepravidelné vzorkování,
- selektivní vzorkování,
- progresivní vzorkování,
- profilování.

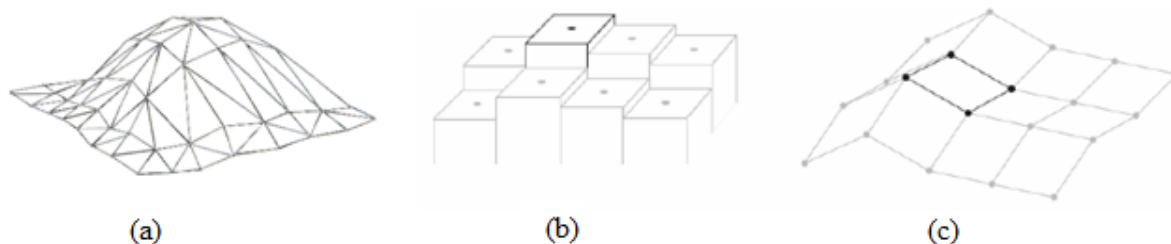


Obr. 5 Postupy při vzorkování (zdroj: <http://www.spatial.maine.edu/~beard/Lectures510/digital%20terrain%20models%2009.pdf>).

Na přesnost výstupního modelu má vliv rozmístění vstupních bodů, jejichž výběr se provádí různými metodami vzorkování. V diplomové práci bylo použito pravidelné vzorkování v různých intervalech, kde data vytváří pravidelnou síť a kvůli potřebě zaznamenat všechny nespojitosti svahů a změny v topografii dochází k nadměrnému získávání dat. Dalším vzorkováním, které bylo použito, je náhodné vzorkování, které může vytvářet nežádoucí shluky.

3.2.3 Typy datového modelu

Hlavním cílem použití vhodného datového modelu při tvorbě digitálního modelu je získání optimální reprezentace reliéfu v GIS při použití co nejmenšího množství dat. Při nadbytečném množství dat se zvyšuje velikost souboru, náročnost na zpracování, možnost redundance dat atd. Podle typu základní geometrické jednotky máme polyedrický model reliéfu (TIN), rastrový model reliéfu (GRID) a plátový model reliéfu (lattice) (Moore et al. 1991), které jsou na obrázku 6.



Obr. 6 TIN (a), grid (b) a lattice (c) reprezentující terén (zdroj: Wilson a Gallant (2000)).

V závislosti na výškových datech a typu následných analýz byl v diplomové práci pro vyjádření zemského povrchu zvolen *rastrový model reliéfu (GRID)*, který je nejčastěji využívaným datovým modelem. V převážné většině je tvořen čtvercovou maticí bodů, kde každá buňka, kromě svojí pozice (číslo řádku a sloupce), nese hodnotu nadmořské výšky vztaženou ke středu buňky (grid). Jeho značnou výhodou je pravidelná struktura, takže lze snadno provádět statistické analýzy za použití mapové algebry. Grid má také své nevýhody, a to převážně velký objem dat při ukládání, nepřesnosti v hrubém rastru a nadbytečnost údajů na plochách, kde je průběh terénu pravidelný (Voženílek a kol., 2001).

3.2.4 Metody interpolace

Snad nejzávažněji může výsledný digitální model ovlivnit výběr interpolačních metod a nastavením jejich parametrů lze také určit a potvrdit závislost přesnosti právě na vybrané interpolační metodě a jejím nastavení. Do procesu tvorby digitálního modelu reliéfu vstupují izolované výškové body s různou hustotou a rozložením. S cílem nalezení nejvhodnější interpolační metody, která co nejuvěrněji zachytí průběh modelovaného reliéfu. V místech, kde nejsou k dispozici data, dochází k odhadu hodnot a k následnému generování DMR se používají metody prostorové interpolace. Kvalita výsledného modelu je určena rozmístěním a přesností původních bodů a vhodností interpolační metody pro určitý model (Voženílek a kol., 2001).

Každá interpolační metoda má své klady a zápory a každé území má svá charakteristická specifika, proto vybrat plošně nejvhodnější metodu interpolace je obtížné a časově náročné. Z tohoto důvodu bylo v diplomové práci počítáno převážně jen s jednou interpolační metodou a to s nástrojem implementovaném v ArcGIS - Topo To Raster, který kombinuje několik interpolačních metod, aby vytvořil hydrologicky korektní model terénu. V práci byly generovány gridy s různou hustotou a rozmístěním vstupních bodů – náhodně rozmístěné body o různých hustotách, pravidelně rozmístěné body a vrstevnice. Na různém nastavení interpolační metody lze velmi dobře sledovat vliv členitosti reliéfu. Hustota a rozmístění bodů byla zvolena tak, aby byly znatelné rozdíly, které jsou způsobené dostatečně vysokou hustotou vstupních dat nebo naopak klesající hustotou vstupních dat, kdy dochází ke vzniku větších rozdílů mezi jednotlivými DEM.

Topo to raster je metoda kombinující výhody následujících interpolací. Je výpočetně méně náročná než lokální interpolační algoritmy např. IDW, přitom výsledný povrch neztrácí návaznost jako u Spline nebo Kriging. Metoda je založená na programu ANUDEM vyvinutém Michaelem Hutchinsonem (1988, 1989, 1996, 2000, 2011). Aktuálně je přístupná verze ANUDEM 5.3. Algoritmu, se kterým pracuje, výrazně zvyšuje přesnost digitálního výškopisného modelu a to i z velice rozsáhlých datových souborů. Nástroj Topo to Raster odhaduje výškové hodnoty pro rastr, zatímco využívá omezení, která se starají o spojenou odtokovou síť a přesnou reprezentaci hřbetů a toků ze vstupních dat. Pro zlepšení chodu interpolace je možné vstupním datům manuálně přiřazovat jejich význam (ArcGIS Help 10.1).

Voženílek a kol. (2001) uvádí další 4 metody prostorové interpolace – IDW (metoda vážených inverzních vzdáleností), spline, kriging a trend. Ale z důvodu časové náročnosti již byl dále testován jen kriging a to nad datovou sadou uměle generovaných dat.

3.2.5 Morfometrické charakteristiky

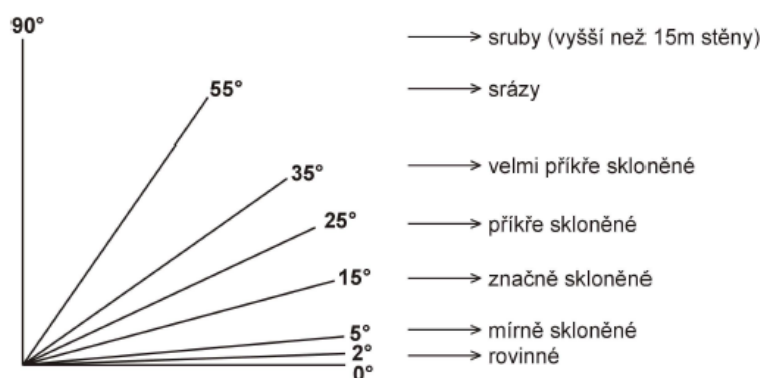
Reliéf se vyznačuje mnoha různými charakteristikami. Wilson a Gallant (2000) dělí morfometrické charakteristiky na primární a sekundární. Primární morfometrické charakteristiky jsou odvozeny pomocí prvních či druhých derivací povrchu (nadmořské výšky), tedy přímo z digitálních výškových modelů a patří mezi ně např. sklon svahu, orientace svahu, horizontální a vertikální křivost. Sekundární morfometrické charakteristiky zahrnují kombinaci dvou nebo více primárních charakteristik nebo kombinaci primární charakteristiky s fyzikálními indexy.

V této práci budou popsány pouze ty analýzy, které nejčastěji slouží k popisu tvarů reliéfu a lze je využít při hodnocení kvality digitálních výškových modelů (Carlisle, 2005), což jsou primární morfometrické parametry odvozené z nadmořské výšky.

- **Sklon svahu**

Sklon svahu (angl. slope) je důležitý morfometrický ukazatel, který určuje míru intenzity geomorfologických projevů. Představuje míru změny nadmořské výšky ve směru nejprudšího poklesu (maximálního spádu). Vychází z první derivace povrchu a odpovídá úhlu, který svírá v libovolném bodě tečna k rovině reliéfu s rovinou horizontální (obr. 7). Může být určován ve stupních, procentech nebo radiánech (Krcho, 1990).

V programu ArcGIS 10.1 se jedná o fokální funkci, která počítá hodnotu sklonu svahu pro každý pixel z hodnot nadmořských výšek okolních devíti pixelů. Nižším hodnotám sklonu svahu odpovídá plošší terén, zatímco vyšší hodnoty znázorňují strmější povrch. Sklon svahu lze tak částečně považovat za indikátory chyb.



Obr. 7 Rozdělení svahů podle míry sklonu

(zdroj: http://geography.upol.cz/soubory/lide/smolova/GMOR/GMOR_Terminologie.pdf).

- **Orientace reliéfu**

Orientace reliéfu (angl. aspect) představuje orientaci povrchu vůči světovým stranám. Je založena na první derivaci povrchu stejně jako sklon svahu. Určuje směr maximálního poklesu výškové hodnoty z každé buňky do buňky sousední. Vyjadřuje se ve stupních v rozpětí 0 - 360° ve směru hodinových ručiček od severu. Hodnota 90° určuje východ, 180° jih, 270° západ, 360° sever. V souladu s geomorfologickou terminologií (rovina je část reliéfu se sklonem do 2°) se orientace reliéfu používá pouze pro území se sklonem větším než 2°. S menším sklonem se pak orientace stává méně významnou až matematicky nedefinovanou (Voženílek a kol., 2001).

Reklasifikace je jednou ze základních operací používaných v GIS, při které dochází k přerozdělování či slučování hodnot ve vstupní rastrové vrstvě do nově vznikající vrstvy. Proto bylo nutné vrstvu orientace reliéfu, která vyjadřuje jednotlivou tematiku, reklasifikovat do klasifikací běžně užívaných v geografii. Orientaci do osmidílné směrové růžice a to do směrů sever, severovýchod, východ, jihovýchod, jih, jihozápad, západ a severozápad. V diplomové práci bylo zjištěno, že orientace svahů příliš nezávisí na typu reliéfu a tím pádem nemá vliv na hodnoty chyb.

- **Křivosti reliéfu**

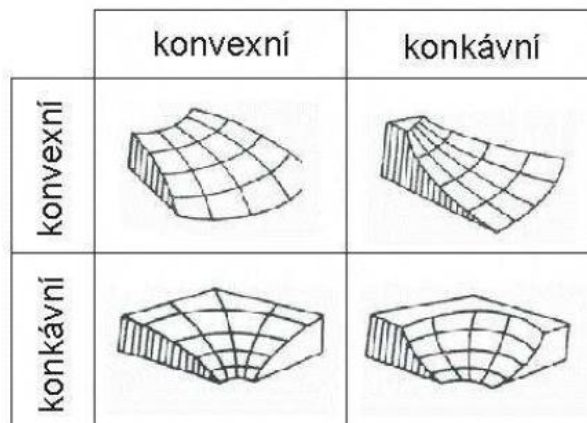
V diplomové práci lze hodnoty křivosti alespoň částečně považovat za indikátory výskytu chyb, to dokazují výrazně vyšší hodnoty regresních koeficientů β , které jsou uvedeny dále. Mezi nejčastější typy křivosti patří planární (horizontální), profilová (vertikální) a celková křivost reliéfu. Další typy křivosti popisuje Jeff Jenness (2013) ve svém manuálu DEM Surface Tools for ArcGIS.

- Planární (horizontální) křivost reliéfu

Planární křivost reliéfu (angl. plan curvature) vychází z druhé derivace povrchu a představuje zakřivení ve směru vrstevnice, udává míru změny orientace. Pozitivní křivost výstupního DMR indikuje konvexní (vypuklý) tvar v dané buňce, negativní křivost konkávní (vyduté), obr. 8. Nulová hodnota je přiřazena k povrchu plochému (Voženílek a kol., 2001).

- Profilová (vertikální) křivost reliéfu

Profilová křivost reliéfu (angl. profile curvature) udává míru změny sklonu povrchu. Představuje zakřivení ve směru spádové křivky. Stejně jako u planární křivosti označují kladné hodnoty konvexní tvary (horské hřbety), záporné konkávní tvary (údolí), nulové plochý povrch (Voženílek a kol., 2001).



Obr. 8 Znázornění kombinací konvexních a konkávních tvarů reliéfu
(zdroj: http://geography.upol.cz/soubory/lide/smolova/GMOR/GMOR_Terminologie.pdf).

- Celková křivost reliéfu

Celková křivost reliéfu vyjadřuje vlastní křivost povrchu, ne křivost linie napříč povrchem v nějakém směru. Může být pozitivní nebo negativní, s nulovým zakřivením znamená, že povrch je buď plochý nebo je konvexnost v jednom směru vyvážena konkavitou v druhém směru, jako např. v sedle (Janness, 2013).

- **Převýšení**

Se sklonem svahu souvisí i převýšení. Z principu funkce algoritmu „slope“, který využívá rozdílů nadmořských výšek mezi dvěma sousedními pixely, vyplývá, že čím větší bude rozdíl těchto výšek, tím větší bude i sklon. Převýšení lze definovat jako rozdíl nadmořských výšek dvou bodů. U převýšení platí, že čím menší převýšení, tím menší nepřesnost a naopak.

3.3 Metody hodnocení přesnosti

Podstatnou částí práce je hodnocení výškových modelů jak pomocí globálních neprostorových, tak lokálních prostorových charakteristik. Značnou výhodou globálních neprostorových charakteristik je jednoduchý výpočet a interpretace. Také znalost rozložení chyb v prostoru je velmi důležitá, protože velikost chyb v DEM je závislá na charakteru terénu a je prostorově proměnlivá. V diplomové práci proto budou pro hodnocení přesnosti DEM použity jak neprostorové tak prostorové metody.

Neprostorové (globální) metody pro hodnocení přesnosti

- **Střední kvadratická chyba**

Nejčastěji používanou metodou je výpočet střední kvadratické chyby (RMSE – root mean square error), která vyjadřuje míru, s jakou se interpolovaná hodnota liší od skutečné hodnoty. Čím větší hodnota, tím větší rozptyl mezi dvěma datovými sadami. Vzorec pro výpočet RMSE je následující (Wood, 1996):

$$RMSE_z = \sqrt{\frac{1}{n} \sum_{i=1}^n (Z_{di} - Z_{ri})^2} \quad (2)$$

kde: Z_{di} je i -tá hodnota nadmořské výšky z povrchu DEM
 Z_{ri} je korespondující nadmořská výška z referenčních dat
 n je počet kontrolovaných bodů

Větší hodnota RMSE odpovídá většímu rozptylu mezi dvěma datovými sadami, ideální hodnota by neměla přesáhnout polovinu hodnoty intervalu původních vrstevnic (Svobodová, 2008).

- **Celková absolutní chyba**

Celková absolutní chyba (AE – absolute error) vyjadřuje skutečnou velikost (sumu) všech odchylek od referenčních dat v kladném i záporném směru (Svobodová a kol., 2009). Matematicky je absolutní chyba vyjádřena jako:

$$AE = \sum_{i=1}^n |Z_{di} - Z_{ri}| \quad (3)$$

kde: Z_{di} je i -tá hodnota nadmořské výšky z povrchu DEM
 Z_{ri} je korespondující nadmořská výška z referenčních dat

K základnímu popisu přesnosti digitálních výškových modelů můžeme dále zařadit např. střední chybu, minimální či maximální chybu (Erdogan, 2009). Dále také rozsah, rozptyl, směrodatnou odchylku a průměr z hodnot výsledných DEM (Fencík a Vajsábllová, 2006). V práci se dále zaměříme na výpočet směrodatné odchylky, průměru a minimální a maximální chyby.

- **Směrodatná odchylka**

Směrodatná odchylka (SD – Standard Deviation) určuje, jak moc jsou hodnoty rozptýleny či odchýleny od průměru hodnot. Hodnoty směrodatné odchylky a střední kvadratické chyby by měly být stejné nebo aspoň podobné. Jestliže je $SD > RMSE$, pak interpolační metoda nadhodnocuje primární proměnné a pokud je $SD < RMSE$, tak jsou primární proměnné podhodnoceny (Li a Heap, 2008).

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (4)$$

kde: x_i je i -tá hodnota nadmořské výšky z povrchu DEM

N je počet kontrolovaných bodů

- **Průměr**

Aritmetický průměr, nebo často též jen průměr, je součet všech hodnot vydělených jejich počtem.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (5)$$

kde: n je počet kontrolovaných bodů

- **Minimální a maximální chyba**

Minimální chyba představuje nejnižší hodnotu ze všech vstupních parametrů. Oproti tomu maximální chyba představuje hodnotu nejvyšší.

Prostorové (lokální) metody pro hodnocení přesnosti

- **LISA**

Lokální indikátor prostorového shlukování (LISA - local indicators of spatial association) se stal v současné době standardním nástrojem pro lokální analýzu. Nejčastěji je využívána jako ekvivalent Moranova I kritéria, které měří prostorovou autokorelaci na základě lokalizace prvků a jejich atributových hodnot. Dále může být také využíván k identifikaci zdrojů prostorové nestacionarity, k identifikaci prostorových shluků hodnot s podobnou velikostí či k existenci výrazných prostorových odchylek (Anselin, 1995). Pokud je určena hladina významnosti, tak výsledky analýzy prostorové autokorelace lze zobrazit v Moranově diagramu (viz obr. 9) nebo také v mapě. Přínosem statistické analýzy LISA je statistické zhodnocení tvorby prostorových shluků včetně zobrazení oblastí s nadprůměrnými a naopak podprůměrnými hodnotami sledovaného jevu v prostoru (Spurná, 2008).

vážená hodnota proměnné v blízkých jednotkách	nízká - vysoká <i>negativní prostorová autokorelace</i>	vysoká - vysoká <i>pozitivní prostorová autokorelace</i>
	nízká - nízká <i>pozitivní prostorová autokorelace</i>	vysoká - nízká <i>negativní prostorová autokorelace</i>
	hodnota proměnné v prostorové jednotce	

Obr. 9 Moranův diagram
(zdroj: Svobodová, 2011).

- Mezi prostorové metody řadíme i geograficky váženou regresi (*GWR*), která je podrobněji popsána níže v kapitole 3.4.

3.4 Hodnocení vztahů mezi parametry

K výzkumu závislostí mezi dvěma či více proměnnými slouží korelační a regresní analýza. Cílem korelace je ověřit existenci či neexistenci vztahu mezi proměnnými a také těsnot závislosti. Pokud se závislost prokázala, bylo možné vypočítat regresní analýzu. Korelační a regresní analýza byla aplikována na hodnoty chyb a hodnoty zkoumaných parametrů.

- **Korelační analýza**

Korelační analýza se zabývá nalezením závislosti mezi proměnnými a nemá definováno, která z proměnných je závislá (vysvětlovaná) a která nezávislá (vysvětlující) proměnná.

Mezi nejznámější výpočetní korelační koeficienty mezi proměnnými patří Pearsonův a Spearmanův korelační koeficient. Pearsonův korelační koeficient vyžaduje, aby parametry měly normální rozdělení, pokud nemají, dochází k chybným výsledkům. Z tohoto důvodu, kdy u některých z testovaných proměnných (např. hodnoty orientace a křivosti) nebyla splněna podmínka normálního rozdělení, byl zvolen jednoduchý typ korelace založený na výpočtu *Spearmanova korelačního koeficientu*, který popisuje vztah mezi proměnnými pomocí monotónní funkce. Hodnoty koeficientu jsou v intervalu $<-1,1>$, kdy kladné hodnoty znamenají pozitivní závislost (rostoucí monotónní funkce), záporné hodnoty negativní závislost (klesající monotónní funkce) a nulová hodnota koeficientu znamená nezávislost mezi proměnnými. Ke zhodnocení těsnoti závislosti byla využita klasifikace uvedená v tabulce 4 (Anděl, 1993). Vzorec pro výpočet korelace je následující:

$$r_s = 1 - \frac{6 \sum_{i=1}^n (i_x - i_y)^2}{n \cdot (n^2 - 1)} \quad (6)$$

kde: i_x, i_y jsou hodnoty dvojice prvků se stejným pořadím
 n je počet sledovaných dvojic

Tab. 4 Těsnost a typ závislosti mezi proměnnými (Anděl, 1993)

Absolutní hodnota r_s	Těsnost závislosti	Typ závislosti
0,0	Nulová	Nezávislost
0,0–0,3	Nízká	Volná závislost
0,3–0,5	Mírná	
0,5–0,7	Význačná	
0,7–0,9	Velká	
0,9–0,99	Velmi vysoká	
1,00	Pevná funkční	Pevná závislost

- **Regresní analýza**

Základním cílem regresní analýzy je nalezení regresní funkce, která co nejlépe vystihuje zkoumané vztahy mezi závisle proměnnou a jednou či více nezávisle proměnnými a vyjádření této funkce matematicky. V regresní analýze je určeno, která z proměnných je vysvětlovaná (závislá) a vysvětlující (nezávislá) a umožňuje předpovídat změnu hodnoty funkční závislosti.

V diplomové práci byl ke stanovení vztahu mezi hodnotami chyb a hodnotami zkoumaných parametrů použit jak lineární model OLS regrese, tak i geograficky vážená regrese, které jsou popsány níže.

Vícenásobná lineární OLS regrese – metoda nejmenších čtverců

K určení vztahu mezi hodnotami chyb a hodnotami či změnami hodnot zkoumaných parametrů odvozenými z DEM byl použit vícenásobný lineární model OLS regrese – metoda nejmenších čtverců. Tato metoda je založena na principu minimalizace součtu čtverců odchylek závisle proměnné od hodnot nezávisle proměnné. Výsledky u OLS jsou nespolehlivé, pokud jsou dvě nebo více proměnných podobné. Více o interpretaci OLS regrese viz Anděl (1993). Vzorec pro výpočet vícenásobné lineární OLS regrese je následující:

$$y = \alpha + \beta_1 x_1 + \dots + \beta_n x_n + e \quad (7)$$

kde: y je závisle proměnná, e je hodnota chyby

α je prostý člen (intercept)

β_1 až β_n jsou regresní koeficienty jednotlivých nezávisle proměnných

x_1 až x_n jsou nezávisle proměnné, zde hodnota nadmořské výšky, orientace svahů, sklonu svahů, celkové křivosti, planární křivosti a profilové křivosti...
 e je residuální odchylka

Geograficky vážená regrese - GWR

Další snahou o zlepšení regresního modelu bylo zahrnutí i geograficky vážená regrese. Tato metoda patří mezi relativně nové kvantitativní techniky prostorové analýzy dat, byla vyvinuta v 90. letech minulého století. Je to metoda, která odstraňuje nevýhody globálních modelů tím, že pro prostorově definovaná data stanovuje lokální hodnoty parametrů modelu a je založena na principu, že pro každý bod analyzovaného území lze vypočítat lokální hodnoty parametrů na základě určitého počtu okolních bodů. U globálních prostorových statistik popisují studované území jen jedním vztahem a o měnících se vztazích v prostoru nic nevypovídají. Cílem GWR je tak aspoň částečné řešení problému prostorové nestacionarity. Tato metoda není vhodná pro malé datové soubory a staví regresní rovnici pro každý prvek v datovém souboru. V globálních regresních modelech, jako je například OLS, jsou výsledky nespolehlivé, když jsou dvě nebo více proměnných podobné (Drápela, 2006). Výsledkem analýzy pomocí GWR není jedna regresní rovnice, ale soubor regresních koeficientů pro každý regresní bod zvlášť včetně dalších lokálních standardních regresních diagnostik. GWR je vypočítána jako:

$$y(u, v) = \beta_0(u, v) + \beta_1(u, v)x_1 + \varepsilon(u, v) \quad (8)$$

kde: y je závislá proměnná

x_1 je nezávislá proměnná

β_0 a β_1 jsou odhadované parametry

ε je náhodná chyba, u které se předpokládá, že má normální rozdělení

u, v jsou souřadnice bodu

Při zpestřování regresních modelů lze využít i prostorovou autokorelaci (Moranovo I kritérium), která je popsána v kapitole 3.3. Při analýze prostorových dat u regresních analýz se často používá metoda regresních reziduí, jejíž cílem je kartografické vynesení rozdílů mezi teoretickými hodnotami získanými z regresního modelu a empirickými hodnotami závisle proměnné. Analýza rozmístění regresních reziduí umožňuje prostorovou verifikaci závislostí a vylepšení regresních modelů. Grafická analýza regresních reziduí (kladných a záporných hodnot) určí, zda lze zkoumaný vztah považovat za stacionární. V případě zjištění výraznějších shluků regresních reziduí lze očekávat neplatnost globálního vztahu, a proto je nutné zabývat se prostorovou heterogenitou a nestacionaritou, tedy problémem, že se charakter určitých jevů a vzájemných vztahů mezi nimi odlišuje v prostoru. Analýza rozmístění regresních reziduí, doplněná výpočtem koeficientů prostorové autokorelace, je často prvním krokem vedoucím k uvažování o prostorové nestacionaritě a nutnosti použití novějších metod prostorové analýzy dat a to např. metodu geograficky vážené regrese (Spurná, 2008).

4 TVORBA DEM A ODVOZENÍ ZKOUMANÝCH PARAMETRŮ

V této kapitole je popsána příprava a sestavení DEM, použití interpolační metody a základní analýza vstupních dat. V diplomové práci se pracuje s umělými náhodně generovanými daty a s reálnými daty z leteckého laserového skenování. První část práce je zaměřena na zpracování dat z matematických modelů a pro ověření a zhodnocení výsledků se v práci využívají laserová data, která jsou mezi sebou v závěru práce porovnána. Zpracování dat pro každé území bylo provedeno převážně v prostředí ArcGIS Desktop a statistickém programu R.

4.1 Získání a úprava vstupních dat

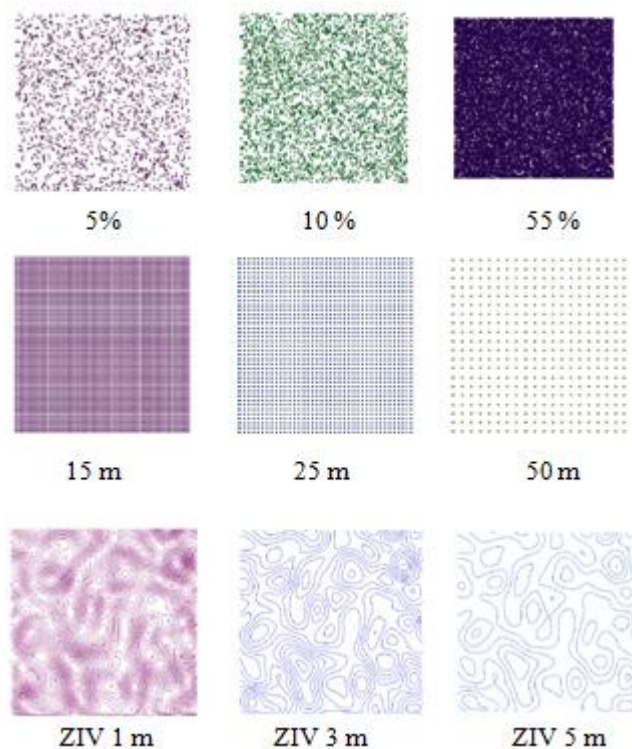
- **Umělá, náhodně generovaná data**

Jako testovací datová sada byla použita umělá data náhodně generovaná v programu R, kde bylo vygenerováno 40 000 bodů v pravidelné mřížce. Každý bod obsahoval souřadnici x a y a výšku z . Následné operace byly provedeny v prostředí ArcGIS Desktop, kde bylo nutné, dle cíle práce, vytvořit různé rozmístění vstupních bodů (obr. 10). Náhodně rozmístěné body o různých hustotách byly tvořeny pomocí funkce *Geostatistical analyst extension/Create Subsets* a vybrán náhodný výběr bodů vstupující do výpočtu DEM, dle vlastního stanoveného procentuálního poměru, a to 5 %, 10 % a 55 %. Pravidelně rozmístěná síť bodů mezi sebou byla vytvořena ve vzdálenosti 15 m, 25 m a 50 m pomocí volně dostupné extenze *Repeating Shapes*. A pomocí funkce *Spatial Analys/Contour* byly vytvořeny vrstevnice v základním intervalu 1 m, 3 m, 5 m. Hustota a rozmístění bodů byla zvolena tak, aby byly znatelné rozdíly, které jsou způsobené dostatečně vysokou hustotou vstupních dat či naopak klesající hustotou vstupních dat, kdy dochází ke vzniku větších rozdílů mezi jednotlivými DEM.

- **Data z leteckého laserového skenování**

Data leteckého laserového skenování byla získána ve formátu *.las. Zpracována byla pomocí nástrojů z nabídky *Spatial Analyst Tools* v prostředí ArcGIS společnosti ESRI. Nejprve byl využit nástroj *Point File Information*, který zjišťuje základní statistické charakteristiky LAS souboru. Výstupem je Shapefile (SHP soubor), jehož atributová tabulka udává celkový počet zaměřených bodů 8 354 301, dále také střední vzdálenost bodů a minimální a maximální hodnotu souřadnice Z . Tato atributová tabulka je na obrázku č. 11. Dalším krokem je použití nástroje *LAS to Multipoint*, který dokáže načíst LAS soubory do geodatabáze. Zde je třeba vložit vybraná čísla kategorií dle standardu ASPRS, která jsou v práci uvedena v kapitole 2.1 v tabulce č. 1. Při výběru dat pro tvorbu DEM byla vybrána pouze kategorie 2 - ground, ve které jsou obsaženy body odražené od zemského povrchu. Výstupní třída prvků obsahuje 1 436 záznamů. V atributové tabulce se nachází pouze tři atributy, kde prvním z nich je OID - identifikační číslo objektu. Druhý je atribut Shape, jehož obsahem je „Multipoint Z“.

A třetím z atributů je PointCount, kde je u každého záznamu uveden počet bodů, které jsou uloženy pod jedním záznamem v geodatabázi, což umožňuje velkou kompresi třídy a je pro data leteckého laserového skenování ideální. V běžné bodové vrstvě je pod jedním záznamem uložen pouze jeden bod (Lidar Analysis in ArcGIS, 2010).



Obr. 10 Rozmístění vstupních bodů (náhodné, pravidelné, vrstevnice) z uměle generovaných dat.

FID	Shape *	FileName	Pt_Count	Pt Spacing	Z Min	Z Max
0	Polygon	Zlin_0-5.las	1982593	0,480751	403,75	521,93
1	Polygon	Zlin_0-6.las	3735115	0,416907	309,96	1072,24
2	Polygon	Zlin_1-5.las	846265	0,486943	441,37	520,61
3	Polygon	Zlin_1-6.las	1790328	0,395168	411,08	605,48

Obr. 11 Atributová tabulka SHP souboru vytvořeného nástrojem *Point File Information*.

4.2 Tvorba DEM

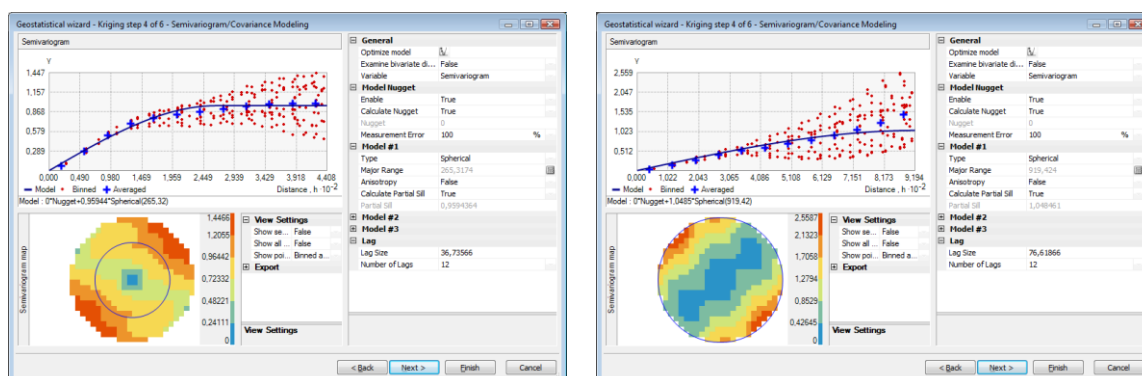
Pro tvorbu DEM byla použita interpolační metoda implementována v prostředí ArcGIS Desktop Topo to Raster, která pomocí funkce *Spatial Analyst Tools/Interpolation/Topo to Raster* kombinuje několik interpolačních metod, je výpočetně méně náročná než lokální interpolační algoritmy např. IDW, přitom výsledný povrch neztrácí návaznost jako u Spline nebo Kriging. Následně bylo vygenerováno 9 gridů z uměle generovaných dat a 9 gridů z dat leteckého laserového skenování (celkem tedy 18 gridů) na základě stejného rozmístění vstupních dat (tab. 3) a s rozlišením pixlu 5 m. Velikost území z uměle generovaných dat je 1 x 1 km a z dat leteckého laserového skenování je 1,1 x 1,4 km. Jako bezchybný (originální) model je brán model ze 100 % původních bodů, tj. 40 000 bodů, kde byla použita interpolační metoda Topo to Raster a který slouží pro následné hodnocení přesnosti DEM.

4.3 Základní analýza dat

Před použitím statistických metod pro hodnocení DEM je třeba provést ověření základních charakteristik vstupních dat a to závislosti (autokorelace) a normality rozdělení. Do analýzy vstupoval 100% DEM, které slouží pro následné hodnocení přesnosti. Dle Erdogana (2009) by nemělo význam data dále analyzovat, pokud by nevykazovala závislost. Cílem je poukázat na případné porušení některého z požadavků, které mohou ovlivnit výsledné DEM.

- **Závislost (autokorelace)**

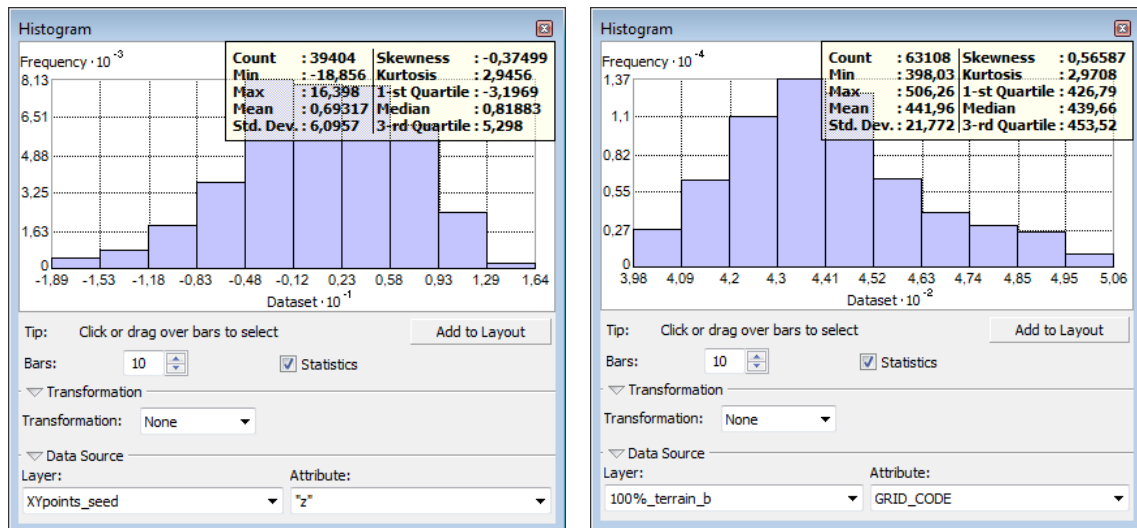
Prostorová závislost dat je nejčastěji modelována pomocí semivariogramu. Pokud by mračno bodů v semivariogramu nevykazovalo závislost, bylo by nemožné odhadovat neznámé hodnoty. Při pohledu na semivariogram ze vstupních dat matematicky generovaného modelu a z dat leteckého laserového skenování lze říci, že data v modelovaném území vykazují závislost (obr. 12).



Obr. 12 Příklad semivariogramu z matematického modelu (vlevo) a laser scanu (vpravo).

- **Normalita rozdělení**

Normalitu rozdělení vstupních dat lze hodnotit např. pomocí míry šikmosti a špičatosti nebo graficky např. pomocí histogramu. Šikmost udává, zda jsou hodnoty rozloženy souměrně, zda je rozdělení hodnot zešikmeno či zda je symetrické. V případě normality je hustota pravděpodobnosti vzhledem ke střední hodnotě symetrická a koeficient šikmosti je nulový (Anděl 1993). Špičatost vyjadřuje, že čím je rozdělení špičatější, tím víc jsou hodnoty soustředěny kolem daného středu. Za předpokladu normality je koeficient špičatosti k_4 roven 3 (Anděl 1993). Tvar histogramu (obr. 13) by v ideálním případě měl odpovídat gaussovu rozdělení, tj. nejvyšší četnosti dosahují hodnoty okolo středu. Pro tvorbu histogramu byl využit nástroj *Geostatistical Analyst extension/Explore data/Histogram*, který zároveň umožňuje i výpočet hodnot šikmosti a špičatosti charakterizující normalitu rozdělení.



Obr. 13 Ukázka tvaru histogramu matematického modelu (vlevo) a laser scanu (vpravo).

Z tabulky 5 vyplývá, že hodnoty špičatosti u obou území se blíží ideální hodnotě 3. Hodnoty šikmosti $\pm 0,4$ společně s vizuální kontrolou histogramu lze považovat za přijatelné a nezpůsobující výrazné odchylení od teoretického normálního rozdělení.

Tab. 5 Koeficient šikmosti (k_3) a špičatosti (k_4)

Území	k_3	k_4
Matematicky generovaný model	-0,37	2,95
Laser scan	0,57	2,97

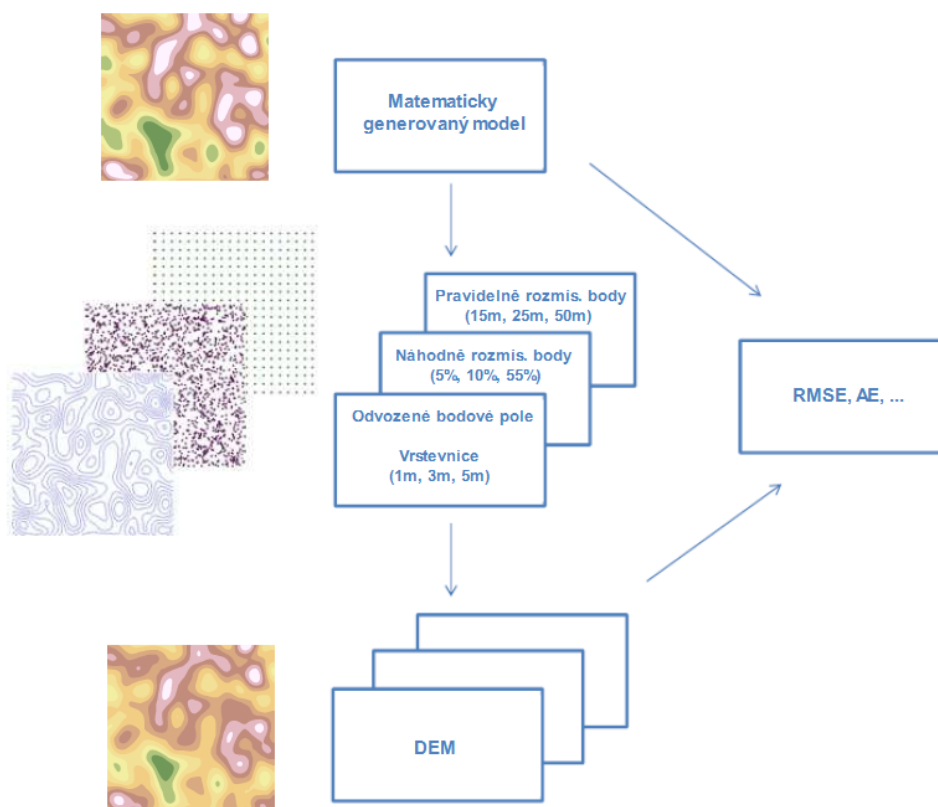
4.4 Odvození zkoumaných parametrů

Po vytvoření DEM (z uměle generovaných dat a dat z leteckého laserového skenování) s různým rozmístěním a hustotou vstupních bodů z nich bylo třeba vypočítat hodnoty zkoumaných parametrů, a to nadmořské výšky, převýšení, vzdálenosti ke vstupním bodům (u vrstevnic k vertexům), sklonu svahu, orientace svahu, orientace reklasifikované, celkové, planární a profilové křivosti, dále extremity minima, maxima a průměru a také parametry získané pronásobením základních morfometrických parametrů sklonu svahu \times profilové křivosti a orientace svahu \times planární křivosti. Celkem vzniklo 252 gridů. Všechny výpočty proběhly v prostředí ArcGIS Desktop. Výpočet sklonu byl vypočítán pomocí nástroje *Slope* z nabídky *Spatial Analyst Tools*, orientace pomocí nástroje *Aspect* a reklasifikovaná orientace byla pomocí funkce *Reclassify* klasifikována do osmi kategorií (sever, severovýchod, východ, jihovýchod, jih, jihozápad, západ a severozápad). Všechny tři křivosti byly vypočítány pomocí funkce *Curvature*. Další zkoumané parametry převýšení a extremity byly počítány pomocí funkce *Focal Statistics* z nabídky *Spatial Analyst Tools* a vzdálenost pomocí nástroje *Near* z nabídky *Analysis Tools*. Výsledné rastry (matice hodnot) následně sloužily k výpočtu korelační a regresní analýzy. Výpočty jsou popsány v kapitole 3.2.5.

5 HODNOCENÍ KVALITY DEM ZE VSTUPNÍCH DAT

Jedním z cílů diplomové práce je hodnocení kvality DEM ze vstupních dat. V této kapitole se zaměříme na hodnocení DEM z matematického modelu, kde jsou použity jak globální (neprostorové), tak lokální (prostorové) ukazatele. Výhodou použití globálních charakteristik, které jsou blíže popsány v kapitole 3.3, je jejich snadný výpočet a jednoznačná interpretace. Díky tomu mohou být DEM seřazeny od nejpřesnějších po méně přesné. Schéma postupu zpracování je vidět na obrázku č. 14.

- **Globální (neprostorové) charakteristiky z matematického modelu**



Obr. 14 Schéma výpočtu globálních charakteristik.

Pro sestavení váženého pořadí DEM jednotlivých modelových území byly zvoleny následující globální charakteristiky: střední kvadratická chyba (RMSE), celková absolutní chyba (AE), minimální a maximální chyba, průměr a směrodatná odchylka. Celkové vážené pořadí DEM daného území bylo získáno součtem dílčích vážených pořadí, kde hodnoty vah byly voleny experimentálně a při použití odlišných charakteristik je možné je modifikovat. Bylo určeno následovně: *vážené pořadí DEM = (pořadí RMSE · 2/8) + (pořadí AE · 2/8) + (pořadí minimální chyby · 1/8) + (pořadí maximální chyby · 1/8) + (pořadí průměru · 1/8) + (pořadí směrodatné odchylky · 1/8)*.

Výsledné vážené pořadí je průměrem pořadí DEM dle jednotlivých charakteristik, kdy důležitost ukazatelů je zohledněna v hodnotě váhy, kterou je dané pořadí násobeno. Pomocí váženého pořadí sice nelze stanovit ostrou hranici mezi kvalitními a nekvalitními DEM, ale lze vybrat nejkvalitnější DEM dle použitých ukazatelů. Podmínkou je tedy vytvoření sady více DEM, které vstupují do hodnocení přesnosti. Z váženého pořadí DEM jasně vyplývá (viz tab. 6), že celkově nejvhodnější metodou se jeví husté rozmístění vrstevnic se ZIV po 1 m, které dosáhlo nejlepšího hodnocení. Naopak rozmístění vrstevnic se ZIV po 5 m, kde s klesající hustotou vstupních dat dochází i ke vzniku větších rozdílů mezi jednotlivými DEM, se jeví jako nejméně vhodné. Vážené pořadí je značně ovlivněno vysokou hustotou vstupních bodů v závislosti na velikosti území, kdy výškové body z vrstevnic se ZIV po 1 m mají 0,065 bodů/m² (viz tab. 7). Je sice známo, že tvorba DEM z vrstevnic není příliš vhodná, protože dochází k zanedbání výrazných terénních prvků jako jsou např. strže, okraje lomů apod., avšak v diplomové práci tento faktor nemá vliv.

Tab. 6 Neprostorové charakteristiky pro hodnocení přesnosti a jejich vážené pořadí

	RMSE	Pořadí RMSE	AE	Pořadí AE	Minimální chyba	Pořadí minimální chyby	Maximální chyba	Pořadí maximální chyby	Průměr	Pořadí průměru	Směrodatná odchylka	Pořadí směrodatné odchylky	Vážené pořadí
Pravidelné vzorkování													
15 m	0,26	4	940,68	5	-1,10	2	5,10	7	0,0245	5	0,26	4	5
25 m	0,20	3	863,54	4	-1,31	3	3,90	4	0,0237	4	0,20	3	3
50 m	0,46	8	1074,34	6	-2,12	6	4,57	5	0,0294	6	0,46	8	6
Náhodné vzorkování (hustota)													
5%	0,46	7	1746,16	8	-4,35	9	9,78	9	0,0437	8	0,45	7	8
10%	0,42	6	1456,85	7	-4,02	8	7,55	8	0,0364	7	0,42	6	7
55%	0,12	2	104,83	2	-2,02	5	2,99	3	-0,0026	2	0,12	2	2
Vrstevnice													
1 m	0,10	1	18,71	1	-0,85	1	0,97	1	0,0005	1	0,10	1	1
3 m	0,31	5	304,47	3	-1,95	4	2,56	2	0,0076	3	0,31	5	4
5 m	0,67	9	1890,19	9	-3,64	7	4,98	6	-0,0473	9	0,67	9	9

Tab. 7 Hustota počtu vstupních bodů/m² u matematického modelu

	Hustota (počet vstupních bodů/m ²)
Prav. rozmís. body	
15 m	0,004
25 m	0,002
50 m	0,000
Náh. rozmís. body	
5%	0,002
10%	0,004
55%	0,022
Vrstevnice	
1 m	0,065
3 m	0,021
5 m	0,013

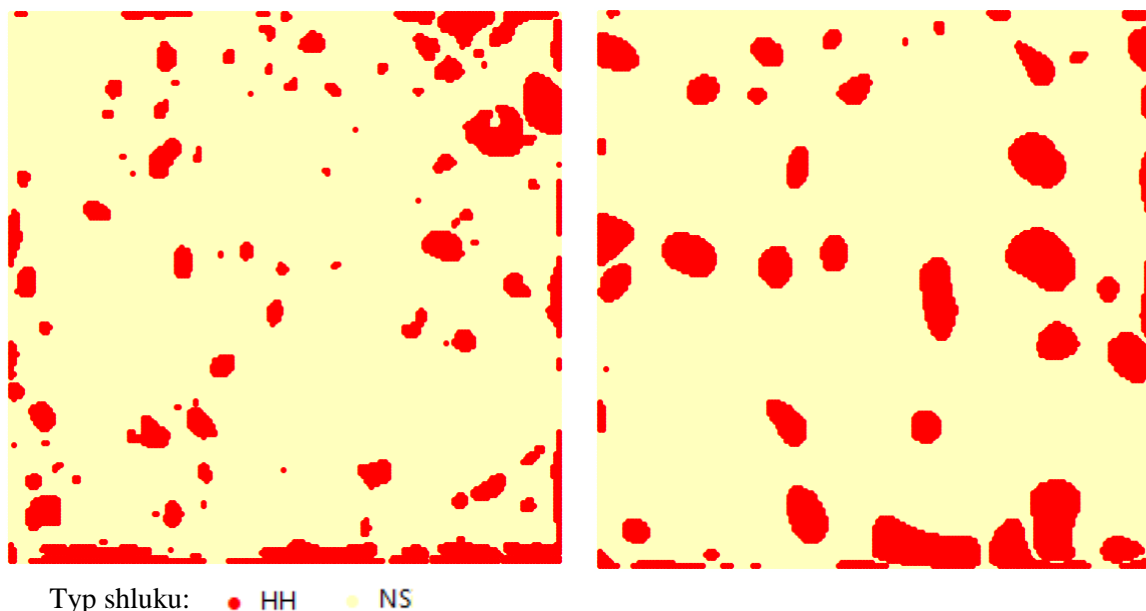
- **Lokální (prostorové) charakteristiky z matematického modelu**

Předcházející metody se zaměřovaly na sledování globálních charakteristik v daném území. Tato část se věnuje prostorovým charakteristikám, které, jak už název napovídá, nelze zobecnit jedním číslem. Jejich hlavním přínosem je zobrazení chyb či nesrovnalostí, proto jsou prostorové charakteristiky posuzovány pomocí vizualizace a to vždy minimálně u dvou DMR. Výše bylo pomocí váženého pořadí zjištěno, že nejvhodnější metodou se jeví husté rozmístění vrstevnic se ZIV po 1 m. Naopak rozmístění vrstevnic se ZIV po 5 m je umístěno ve váženém pořadí na posledním místě, což se jeví jako nejméně vhodné, proto byly tyto dva modely srovnány pomocí lokální shlukové analýzy (LISA).

Program ArcGIS podporuje shlukovou analýzu, kterou je metoda LISA. Ta identifikuje shluky bodů s podobnými hodnotami a shluky s rozdílnými hodnotami, ty jsou kategorizovány jako HH – shluky vysokých hodnot, LL – shluky nízkých hodnot a HL či LH – označují odlehlé hodnoty (outliers). Lisa se nezabývá velikostí hodnot ve shluku. Výpočet lokální shlukové analýzy byl vypočítán v ArcGIS dle nástroje *Spatial Statistic Tools/Mapping Clusters/Cluster and Outlier Analysis (Anselin Local Morans I)*, kam vstupovala bodová vrstva absolutních hodnot chyb, která vznikla odečtením DEM vzniklého interpolací od originálního 100% DEM. Absolutní hodnoty proto, jelikož do analýzy LISA mohou vstupovat pouze nezáporná racionální čísla. Výsledkem metody LISA jsou hodnoty indexu lokálního Moranova I, Z score, p-value a kód typu shluku. Na základě Z score a p-value, které vyjadřují statistickou významnost, je rozhodováno, zda bude potvrzena či zamítnuta nulová hypotéza H_0 . Při potvrzení H_0 jsou hodnoty v území rozmístěny náhodně a nedochází zde k prostorovému shlukování. Při zamítnutí nulové hypotézy je potvrzena příslušnost daného prvku ke shluku vysokých, nízkých či odlehlých hodnot, které se statisticky liší od hodnot očekávaných. Vysoká kladná hodnota Z score a LISA pro daný prvek znamená, že se v jeho okolí nachází prvky s podobnými hodnotami chyb, prvek je tedy součástí shluku. Naopak nízká negativní hodnota Z score a LISA pro daný prvek znamená, že se jedná o statisticky významný (při 0,05) prostorový outlier. Pro prvky s p-value < 0,05 je dále určován typ shluku (Svobodová, 2011).

Při testování statistické významnosti a při výpočtu lokálního Moranova I hraje důležitou roli průměrná hodnota chyb, která je pro soubory hodnot chyb odvozených z kvalitního a nekvalitního DEM často velmi rozdílná. Předpokládáme, že u souborů chyb s menším rozsahem a menší průměrnou hodnotou budou i menší odchylky klasifikovány jako statisticky významné a naopak. V našem případě se nelze zaměřit pouze na prostorovou vizualizaci, ale je třeba se zaměřit i na globální hodnoty ukazatelů. V tabulce 8 je vypočítán rozsah hodnot chyb, průměrná hodnota chyb a podíl pixelů patřící k jednotlivým typům shluků. Na první pohled rozsáhlejší shluky vysokých chyb u výsledků z kvalitnějšího DEM ve skutečnosti zahrnují mnohem menší hodnoty chyb než shluky u méně kvalitních DEM (viz obr. 15). Výsledkem LISA u kvalitnějších DEM z vrstevnic se ZIV po 1 m jsou tedy méně rozsáhlé shluky a u méně kvalitních DEM

(dle váženého pořadí) z vrstevnic se ZIV po 5 m jsou kompaktnější a plošně rozsáhlejší shluky vysokých hodnot chyb. U okrajů jednotlivých DEM dochází k většímu shlukování, což je způsobeno větší chybou interpolace, nedostatkem dat. Z vizuálního hodnocení vychází, že v DEM, označovaném jako kvalitnější, se nachází menší množství statisticky významných hodnot chyb než v DEM označeném jako méně kvalitní. Určení kvalitních a méně kvalitních DEM, vycházející z váženého pořadí, je tedy správné.



Obr. 15 Ukázka výsledků lokální shlukové analýzy chyb (LISA) pro matematicky generovaný model. Vlevo je zobrazen DEM z vrstevnic po 1 m (dle váženého pořadí považovány za kvalitnější) a vpravo DEM z vrstevnic po 5 m.

Pozn.: Typy shluků: HH – shluk vysokých hodnot, NS – prvky, které jsou nositeli statisticky nevýznamných hodnot.

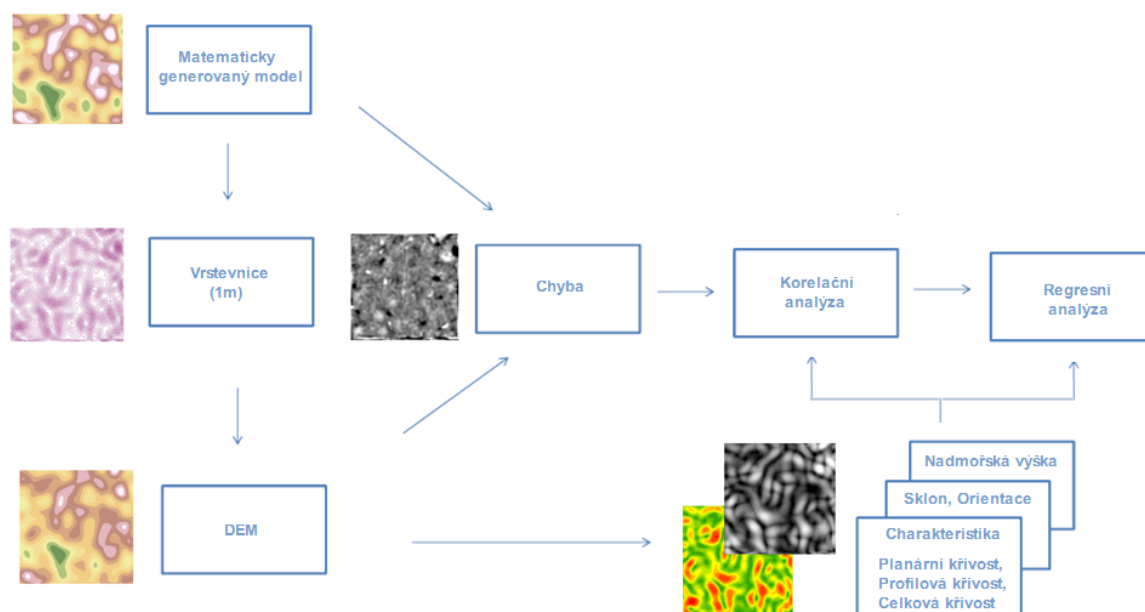
Tab. 8 Tabulka rozsahu, průměru z absolutních hodnot chyb a minimální a maximální hodnoty chyb v jednotlivých typech shluků, kde vrstevnice se ZIV po 1 m jsou dle váženého pořadí považovány za kvalitnější DEM než vrstevnice se ZIV po 5 m

Matematicky generovaný model	Rozsah hodnot chyb [m]	Průměr hodnot chyb [m]	Typ shluku			
			HH		LH	
			min.	max.	min.	max.
vrstevnice 1m	0,97	0,06	0,09	0,97	0,00	0,02
vrstevnice 5m	4,98	0,39	0,86	4,98	-	-

Hlavní význam použití shlukové analýzy v diplomové práci je v lokalizaci vysokých či nízkých hodnot chyb, tedy míst, na která je třeba se více zaměřit při odstraňování chyb a následně je více kriticky zhodnotit. Následně je možné výsledky použít i pro výzkum vztahů mezi výskytem vysokých hodnot chyb v DEM a hodnotami zkoumaných parametrů odvozenými z těchto DEM, protože vymezuje statisticky významné vysoké a nízké hodnoty chyb, popřípadě odlehlé hodnoty.

6 ANALÝZA VZTAHŮ MEZI HODNOTAMI CHYB A HODNOTAMI ZKOUMANÝCH PARAMETRŮ

K výzkumu závislostí mezi proměnnými slouží korelační a regresní analýza. Nejprve byla provedena korelační analýza, kde bylo zjištěno, zda existuje či neexistuje závislost (korelace) mezi proměnnými. Pokud se závislost prokáže, je možné vypočítat regresní analýzu. Schéma postupu zpracování je vidět na obrázku č. 16.



Obr. 16 Schéma výpočtu závislosti mezi hodnotami chyb a hodnotami zkoumaných parametrů.

6.1 Korelační analýza matematického modelu

Cílem korelační analýzy byla snaha o nalezení vztahů mezi hodnotami chyb a hodnotami zkoumaných parametrů. Pro výpočet byl použit neparametrický *Spearmanův pořadový korelační koeficient*, který je podrobněji popsán v kapitole 3.4. Jako podkladový DEM pro výpočet těchto parametrů a následné korelace byl zvolen DEM vytvořený ze ZIV po 1m, který se dle váženého pořadí ukázal pro dané modelové území jako nejkvalitnější. Celý proces byl prováděn na umělých matematických datech na území o velikosti 1 x 1 km a v prostředí R, kde do něj vstupovaly hodnoty z oboru racionálních čísel, protože při použití absolutních hodnot by došlo ke zkreslení modelu.

Prvním krokem bylo nutné zjistit, zda mezi sebou zkoumané parametry korelují a mají mezi sebou závislost a následně použití u regresních analýz. Pokud je mezi dvojicí parametrů příliš silná korelace, pak je jeden z nich tzv. vázaný na druhý a lze jej z analýzy vyloučit. Do regresní analýzy pak vstupoval jen jeden ze dvou parametrů a to ten s vyšší hodnotou korelace s hodnotami chyb. Tabulka 9 ukazuje, že vysoká korelace je mezi sklonem a převýšením, či mezi celkovou křivostí a planární s profilovou křivostí. Vysoká korelace je také u parametrů získaných pronásobením základních

morfometrických parametrů sklonu svahu x profilové křivosti a orientace svahu x planární křivosti s křivostmi a dále mezi nadmořskou výškou a extremitami.

Následujícím krokem bylo zhodnocení intenzity vztahů mezi hodnotami chyb a hodnotami zkoumaných parametrů (viz tab. 10). U většiny případů, dle korelačního koeficientu, byla nalezena nízká nebo mírná závislost ($r_s < 0,5$). Význačná závislost ($r_s \in < 0,5; 0,7$) byla prokázána jen v několika málo případech mezi hodnotami chyb a hodnotami celkové křivosti, planární a profilové. Jelikož při analýze korelace není stanoveno, která z proměnných je vysvětlovaná a která vysvětlující, lze konstatovat, že hodnoty křivostí lze alespoň částečně považovat za indikátory výskytu chyb. Díky prokázané závislosti mezi zkoumanými proměnnými (kromě pěti případů v tab. 10, které mají šedý podklad) je možné vypočítat regresní analýzu.

V rámci snahy o zlepšení intenzity vztahů mezi hodnotami chyb a hodnotami zkoumaných parametrů, byly kromě zkoumaných morfometrických parametrů (nadmořská výška, převýšení, vzdálenost, sklon, orientace, orientace reklasifikovaná, celková křivost, planární křivost a profilová křivost) testovány i další proměnné. Dle Carlisle (Carlisle, 2005) to byly extremity minima, maxima a průměru. Kde *extremity minima* se vypočítají jako rozdíl rozsahu a minima nadmořské výšky v okně 10 x 10 pixlů, *extremity maxima* jako rozdíl rozsahu a maxima nadmořské výšky v okně 10 x 10 pixlů a *extremity průměru* jako rozdíl rozsahu a průměru nadmořské výšky v okně 10 x 10 pixlů. Jak je vidět v tabulce 10, hodnoty korelačních koeficientů jsou velmi nízké. Další snahou o zlepšení bylo pronásobení vybraných základních morfometrických parametrů, a to sklon svahu x profilová křivost a orientace svahu x planární křivost. Což však nevedlo ke zlepšení těsnosti závislosti a ani ke zlepšení regresního modelu, který je popsán v následující kapitole. Důvodem byla snaha o zvýšení variability hodnot proměnných a zdůraznění extrémů daných parametrů.

Dle Erdogana (2009), který našel velmi silnou závislost mezi hodnotami chyb v DEM a hodnotami celkové křivosti (hodnoty korelačního koeficientu $> 0,5$) na území o rozloze pouhých 0,12 km², bylo snahou *zmenšit i zkoumané území*, a to na velikost 200 x 200 m. Výsledkem bylo, že intenzity vztahů u vybraných parametrů nepřinesly žádné výrazné zlepšení, převládala tak nízká a mírná těsnost závislosti (viz tab. 11). Zmenšené území proto nebylo dále v regresní analýze využíváno.

Tab. 11 Hodnoty Spearmanova korelačního koeficientu mezi hodnotami chyb a hodnotami základních zkoumaných parametrů - zmenšené území

Matematický terén 200 x 200 m		Morfometrické parametry							
		nadmořská výška	vzdálenost	sklon	orientace	orientace reklasifikovaná	křivost celková	křivost planární	křivost profilová
<i>Vrstevnice</i>									
1 m	chyby	0,282	-0,036	0,069	0,116	0,061	0,331	0,287	-0,303
3 m	chyby	0,394	-0,145	0,295	-0,045	-0,047	0,412	0,383	-0,341
5 m	chyby	0,637	-0,188	0,419	-0,141	-0,077	0,447	0,397	-0,354
<i>Náhodně rozmístěné body</i>									
5%	chyby	0,493	-0,093	-0,148	0,204	0,276	0,394	0,399	-0,237
10%	chyby	0,385	-0,055	-0,323	0,183	0,368	0,342	0,343	-0,263
55%	chyby	0,434	0,022	-0,397	0,280	0,392	0,409	0,435	-0,228
<i>Pravidelně rozmístěné body</i>									
15 m	chyby	0,379	-0,001	-0,361	0,225	0,262	0,409	0,432	-0,236
25 m	chyby	0,462	-0,027	-0,475	0,375	0,034	0,339	0,377	-0,133
50 m	chyby	0,379	-0,016	-0,011	0,140	0,197	0,458	0,408	-0,228

Těsnost závislosti: nulová nízká mírná význačná

O zlepšení výsledků korelační analýzy se dále pokoušel výpočet pomocí *Pearsonova korelačního koeficientu*, který vyžaduje, aby parametry měly normální rozdělení. Proto byly ty parametry, které normální rozdělení nespĺňují (např. hodnoty orientace a křivosti), odstraněny (viz tab. 12). Těsnost těchto vztahů je však stále pouze nízká nebo mírná ($r_s < 0,5$).

Tab. 12 Hodnoty Pearsonova korelačního koeficientu mezi hodnotami chyb a hodnotami zkoumaných parametrů splňujících podmínku normálního rozdělení hodnot

Matematický terén 1 x 1 km		Zkoumané parametry					
		nadmořská výška	převýšení	sklon	extremita maxima	extremita minima	extremita průměru
<i>Vrstevnice</i>							
1 m	chyby	0,252	-0,544	0,006	-0,189	-0,153	-0,194
3 m	chyby	0,232	-0,255	-0,021	-0,198	-0,173	-0,203
5 m	chyby	0,157	-0,134	0,108	-0,108	-0,053	-0,090
<i>Náhodně rozmístěné body</i>							
5%	chyby	-0,018	0,157	0,035	0,010	0,006	0,002
10%	chyby	-0,081	0,382	0,134	0,056	0,053	0,055
55%	chyby	-0,117	0,328	-0,014	0,076	0,056	0,077
<i>Pravidelně rozmístěné</i>							
15 m	chyby	-0,008	0,290	0,072	-0,003	0,001	0,002
25 m	chyby	0,006	0,199	0,020	-0,005	-0,004	-0,006
50 m	chyby	0,188	-0,128	0,039	-0,151	-0,119	-0,145

Těsnost závislosti: nulová nízká mírná význačná

V neposlední řadě bylo zlepšení intenzity vztahů mezi hodnotami chyb a hodnotami zkoumaných parametrů testováno na změně interpolační metody, kdy byla zvolena běžně dostupná interpolační metoda a to *běžný kriging – semivariogram - sférický*. Což však opět nevedlo ke zlepšení těsnosti závislosti. Byla nalezena převážně nízká nebo mírná závislost ($r_s < 0,5$) a nedošlo ani ke zlepšení regresního modelu (viz příloha 1). Proto bylo v diplomové práci od testování na dalších interpolačních metodách upuštěno.

6.2 Regresní analýza matematického modelu

Základním cílem regresní analýzy je nalezení regresní funkce, která co nejlépe vystihuje zkoumané vztahy a vyjádření této funkce matematicky. Na rozdíl od korelační analýzy je zde jasně určeno, která proměnná je ve vztahu vysvětlována a která vysvětlující. Dle některých autorů (Erdogan, 2010 a Carlisle, 2005) nelze velikost chyb vysvětlit jen jedním parametrem, ale je třeba použít více vysvětlujících parametrů zároveň. Proto byla v diplomové práci využita vícenásobná lineární OLS regrese, poté i geograficky vážená regrese (GWR) a jejich výsledky byly porovnány. Do regresních modelů GWR pak vstupovaly nezávisle proměnné, které byly získány pomocí metody OLS, což zaručuje srovnatelnost výstupů obou metod.

6.2.1 Vícenásobná lineární OLS regrese – metoda nejmenších čtverců

K určení vztahu mezi hodnotami chyb a hodnotami či změnami hodnot zkoumaných parametrů odvozenými z DEM byl použit lineární model OLS regrese. Celý proces byl prováděn na umělých matematických datech na území o velikosti 1 x 1 km a v prostředí ArcGIS Desktop. Do výpočtu vstupovala jediná bodová vrstva obsahující v atributové tabulce hodnoty chyb jako závislá proměnná a jako nezávislé proměnné hodnoty zkoumaných parametrů (celkem 14 parametrů - nadmořská výška, převýšení, vzdálenost ke vstupním bodům (u vrstevnic k vertexům), sklon svahu, orientace svahu, orientace svahu reklasifikovaná, planární, profilová a celková křivost, extremita minima, maxima a průměru a také parametry získané pronásobením základních morfometrických parametrů sklonu svahu × profilové křivosti a orientace svahu × planární křivosti). Pro výpočet byl použit nástroj *Spatial Statistics Tools/Modeling Spatial Relationships/Ordinary Least Squares*.

Hodnota koeficientu determinace Adjusted R^2 regresních modelů je dle výsledků velmi nízká (viz tab. 13). Ve většině případů nepřekračuje hodnotu 0,30. Což znamená, že změny hodnot chyb nelze pomocí hodnot zkoumaných parametrů vysvětlit z více jak 30%. Pouze v jednom případě je hodnota koeficientu determinace vyšší, a to 0,31. Dle hodnot regresních koeficientů se na lepším provedení modelu podílejí výhradně morfometrické parametry vyjadřující křivost reliéfu.

Tab. 13 Výsledné hodnoty prostého členu α , regresních koeficientů a koeficientu determinace Adjusted R^2 získané pomocí vícenásobné lineární OLS regrese

Matematický terén 1 x 1 km	α	Regresní koeficienty nezávisle proměnných														Ad. R^2
		β_1 nadmořská výška	β_2 převýšení	β_3 vzdálenost	β_4 sklon	β_5 orientace	β_6 orientace reklasifikovaná	β_7 křivost celková	β_8 křivost planární	β_9 křivost profilová	β_{10} sklon x profilová křivost	β_{11} orientace x planární křivost	β_{12} extremita maxima	β_{13} extremita minima	β_{14} extremita průměru	
<i>Vrstevnice</i>																
1 m	0,006	0,010	0,025	-0,002	-0,011	0,000	0,000	0,040	-0,016	-0,073	-0,009	0,000	-0,002	-0,013	0,026	0,260
3 m	0,045	0,111	0,050	-0,003	-0,019	0,000	-0,006	0,364	-0,339	0,006	0,094	0,000	0,174	-0,033	-0,033	0,129
5 m	0,069	0,213	0,043	-0,011	0,011	0,000	0,000	0,676	-0,451	0,372	0,001	0,134	0,353	-0,113	-0,028	0,114
<i>Náh. rozmís. body</i>																
5%	0,048	-0,088	0,321	0,005	-0,002	0,000	-0,006	0,254	-0,334	0,538	-0,005	-0,001	0,075	0,095	-0,260	0,128
10%	0,089	-0,153	0,383	0,003	0,013	0,000	0,006	0,145	-0,273	0,379	0,003	0,000	0,023	0,111	-0,288	0,311
55%	0,010	-0,016	-0,009	0,000	0,005	0,000	-0,001	0,010	-0,124	-0,056	0,016	0,000	-0,015	0,013	-0,016	0,203
<i>Prav. rozmís. body</i>																
15 m	0,059	0,014	0,454	-0,004	-0,016	0,000	0,004	0,089	0,063	0,431	-0,007	0,000	0,134	-0,089	-0,031	0,259
25 m	-0,013	0,030	0,221	0,000	-0,007	0,000	0,007	0,041	0,053	0,358	-0,026	0,000	0,093	-0,067	0,006	0,192
50 m	-0,099	0,224	0,283	0,001	-0,019	0,001	0,002	0,378	-0,162	0,721	0,003	0,001	0,331	-0,210	0,104	0,144

vysoký faktor VIF (Variance Inflation Factor) VIF > 7,5

Parametry, které jsou v tabulce 13 značeny fialově, vykazují vysokou multikolinearitu. Ta byla ověřována pomocí výpočtu faktoru VIF (Variance Inflation Factor). Pokud byla některá z vysvětlujících proměnných ohodnocena $VIF > 7,5$, byla z modelu odstraněna. Limitní hodnota je dle doporučení ArcGIS Resource Center. Proces výpočtu každého regresního modelu se tedy několikrát opakoval dle principu postupné zpětné regrese (*angl. backward stepwise regression*). Z důvodu multikolinearity byly z výpočtu odstraněny parametry: nadmořská výška, převýšení, sklon, celková, planární a profilová křivost, a také proměnné extrimita maxima, minima a průměru. Výsledné hodnoty po odstranění parametrů jsou znázorněny v tabulce 14. Po odstranění parametrů, dle vysokého faktoru VIF, dochází ještě ke snížení hodnoty koeficientu determinace Adjusted R^2 , než před jejich odstraněním. To je pravděpodobně způsobeno samotným zkrácením a multikolinearitou při vkládání tolika parametrů. Pouze pro území DEM vytvořené ze ZIV po 1 m, který se dle váženého pořadí ukázal jako nejkvalitnější, je hodnota koeficientu determinace R^2 0,21. Tato hodnota znamená, že velikost hodnot chyb nelze pomocí hodnot zkoumaných parametrů vysvětlit z více jak 21%.

Tab. 14 Výsledné hodnoty prostého členu α , regresních koeficientů a koeficientu determinace Adjusted R^2 po odstranění parametrů na základě vysokého faktoru VIF získané pomocí vícenásobné lineární OLS regrese

Matematický terén 1 x 1 km	α						Ad. R^2
		β_1 vzdálenost	β_2 orientace	β_3 orientace reklasifikovaná	β_4 sklon x profilová křivost	β_5 orientace x planární křivost	
<i>Vrstevnice</i>							
1 m	-0,002	-0,001	0,000	0,001	-0,024	0,000	0,210
3 m	0,032	-0,001	0,000	-0,006	-0,021	0,001	0,034
5 m	0,096	-0,010	0,000	0,001	0,003	-0,014	0,070
<i>Náh. rozmís. body</i>							
5%	-0,015	0,005	0,000	-0,005	0,015	-0,001	0,056
10%	0,014	0,003	0,000	0,007	0,023	-0,001	0,110
55%	0,007	0,000	0,000	-0,002	0,011	0,000	0,169
<i>Prav. rozmís. body</i>							
15 m	-0,013	-0,003	0,000	0,007	0,032	0,000	0,081
25 m	-0,043	0,000	0,000	0,009	0,011	0,000	0,039
50 m	-0,071	0,001	0,000	0,007	-0,011	0,003	0,086

Dalším krokem pro zlepšení regresního modelu bylo odstranění parametrů dle výsledků vzájemných korelací mezi proměnnými (viz tab. 9). Tam, kde se vyskytla význačná až velmi vysoká korelace, vstoupil do následného regresního modelu vždy parametr s vyšší hodnotou korelace s hodnotami chyb (dle tab. 10), z důvodu většího vlivu na předpovídání velikosti chyb. Z výpočtu proto byly odstraněny parametry: nadmořská výška, převýšení, orientace, planární a profilová křivost, extrimita minima, maxima a průměru a také parametry získané pronásobením základních morfometrických parametrů sklonu svahu \times profilové křivosti a orientace svahu \times planární křivosti.

Výsledné hodnoty po odstranění parametrů jsou znázorněny v tabulce 15. Výsledné hodnoty koeficientu determinace R^2 jsou velmi nízké, maximálně do 0,22. Ale oproti předchozímu odstranění parametrů dle vysokého faktoru VIF dochází k mírnému zlepšení koeficientů determinace regresního modelu na základě odstranění parametrů dle výsledků vzájemných korelací mezi proměnnými.

Tab. 15 Výsledné hodnoty prostého členu α , regresních koeficientů a koeficientu determinace Adjusted R^2 po odstranění jednoho z dvojice parametrů, u kterých byla prokázána příliš silná korelace získané pomocí vícenásobné lineární OLS regrese

Matematický terén 1 x 1 km	α					Ad. R^2
		β_1 vzdálenost	β_2 sklon	β_3 orientace reklasifikovaná	β_4 křivost celková	
<i>Vrstevnice</i>						
1 m	0,010	-0,002	-0,002	0,001	0,100	0,221
3 m	0,080	-0,002	-0,006	-0,004	0,192	0,060
5 m	0,074	-0,010	0,004	0,001	0,357	0,071
<i>Náh. rozmís. body</i>						
5%	-0,056	0,005	0,008	-0,001	-0,196	0,062
10%	-0,121	0,003	0,022	0,001	-0,218	0,182
55%	0,007	0,000	0,000	-0,002	-0,075	0,145
<i>Prav. rozmís. body</i>						
15 m	-0,051	-0,003	0,007	0,011	-0,128	0,076
25 m	-0,049	0,000	0,002	0,012	-0,071	0,050
50 m	-0,078	0,001	0,005	0,014	0,244	0,052

Dalším přístupem pro zlepšení regresního modelu bylo odstranění kombinací parametrů dle vysokého faktoru VIF a dle výsledků vzájemných korelací mezi proměnnými (kombinace tabulek 14 a 15). Z výpočtu proto byly (oproti prvnímu modelu) odstraněny parametry: nadmořská výška, převýšení, planární a profilová křivost a také extrémita minima, maxima a průměru. Výsledné hodnoty po odstranění parametrů jsou znázorněny v tabulce 16. Do regresního modelu tedy vstupovaly jako jediná bodová vrstva obsahující v atributové tabulce hodnoty chyb jako závislá proměnná a jako nezávislé proměnné hodnoty 7 zkoumaných parametrů.

Už při výpočtu koeficientů korelace byla mezi hodnotami chyb a hodnotami zkoumaných parametrů u většiny případů nalezena nízká nebo mírná závislost ($r_s < 0,5$). Ani výběr parametrů kombinací tabulek 14 a 15 však nepřinesl očekávané zjištění výrazně silnější závislosti mezi proměnnými, ale oproti předchozímu odstranění parametrů dochází k mírnému zlepšení koeficientů determinace maximálně do 0,24, což znamená, že změny hodnot chyb nelze pomocí hodnot zkoumaných parametrů vysvětlit z více jak 24%. Z celkového pohledu na regresní koeficienty β (viz tab. 16) vyplývá, že nejslabší vazba mezi hodnotami chyb a hodnotami zkoumaných parametrů je u vzdáleností ke vstupním bodům (u vrstevnic k vertexům), sklonu svahů, orientací, reklasifikovaných orientací a také parametrů získaných pronásobením základních

morfometrických parametrů sklonu svahů x profilové křivosti a orientací svahů x planární křivosti. Výrazně vyšší hodnoty regresních koeficientů β jsou u celkové křivosti, které mají vyšší efekt na hodnoty chyb.

Za výsledný regresní model (viz níže), kde bylo dosaženo nejlepších výsledků koeficientů determinace, lze považovat parametry v tabulce 16. Jako nezávislé proměnné vstupovaly parametry vzdálenosti ke vstupním bodům (u vrstevnic k vertexům), sklonu svahu, orientace svahu, orientace svahu reklasifikované, celkové křivosti a parametry získané pronásobením základních morfometrických parametrů sklonu svahu x profilové křivosti a orientace svahu x planární křivosti.

$$\text{chyba} = \alpha + \beta_1 \cdot \text{vzdálenost} + \beta_2 \cdot \text{sklon} + \beta_3 \cdot \text{orientace} + \beta_4 \cdot \text{orientace reklasifikovaná} + \beta_5 \cdot \text{celková křivost} + \beta_6 \cdot (\text{sklon} \cdot \text{profilová křivost}) + \beta_7 \cdot (\text{orientace} \cdot \text{planární křivost})$$

Tab. 16 Výsledné hodnoty prostého členu α , regresních koeficientů a koeficientu determinace Adjusted R^2 po odstranění kombinací parametrů dle vysokého faktoru VIF a dle výsledků vzájemných korelací mezi proměnnými (kombinace tabulek 14 a 15) získané pomocí vícenásobné lineární OLS regrese → výsledný regresní model

Matematický terén 1 x 1 km	α								Ad. R^2
		β_1 vzdálenost	β_2 sklon	β_3 orientace	β_4 orientace reklasifikovaná	β_5 křivost celková	β_6 sklon x profilová křivost	β_7 orientace x planární křivost	
<i>Vrstevnice</i>									
1 m	0,009	-0,002	-0,002	0,000	0,001	0,079	-0,010	0,000	0,237
3 m	0,082	-0,003	-0,007	0,000	-0,005	0,334	0,035	0,000	0,068
5 m	0,078	-0,010	0,003	0,000	0,002	0,547	0,001	0,072	0,085
<i>Náh. rozmís. body</i>									
5%	-0,062	0,005	0,008	0,000	-0,006	-0,171	-0,005	-0,001	0,067
10%	-0,118	0,003	0,022	0,000	0,005	-0,233	0,003	0,001	0,190
55%	0,006	0,000	0,000	0,000	-0,002	-0,005	0,010	0,000	0,169
<i>Prav. rozmís. body</i>									
15 m	-0,056	-0,003	0,007	0,000	0,006	-0,191	0,004	0,001	0,120
25 m	-0,059	0,001	0,002	0,000	0,009	-0,212	-0,023	0,001	0,111
50 m	-0,100	0,001	0,006	0,000	0,006	0,054	-0,001	0,002	0,088

Nutnou součástí hodnocení výsledků regrese je také výpočet testu celkové statistické významnosti (*Joint Wald Statistic*), stacionarity (*Koenker Statistic*) a zkreslení modelu (*Jarque-Bera Statistic*) a zhodnocení prostorové autokorelace reziduálních odchylek (pomocí shlukování). Z výsledků těchto testů bylo zjištěno, že všechny vytvořené modely lze považovat za statisticky významné, avšak nestacionární a zkreslené (ve všech případech byla $p\text{-value} < 0,05$). Z důvodu nestacionarity modelu je tedy vhodnější použít prostorový typ regrese – geograficky váženou regresi (GWR), která zohledňuje nestacionární změny hodnot proměnných v prostoru.

Dalším z výsledků OLS regresní analýzy v ArcGIS je i vrstva s hodnotami reziduí. Reziduum můžeme chápat jako velikost chyby, které se v příslušném bodě při odhadu dopouštíme. Obsahuje dva typy reziduí a to:

- *Klasická rezidua* e_i , kde jsou odchylky skutečných hodnot y_i od modelem odhadnutých hodnot \hat{y}_i , tedy $e_i = y_i - \hat{y}_i$.
- *Normovaná rezidua* e_{Ni} , kde rezidua modelu mají normální rozdělení se střední hodnotou 0 a rozptylem 1. Normalizace reziduí se provádí vydělením hodnoty rezidua směrodatnou odchylkou reziduí $e_{Ni} = e_i/s_{ei}$.

Zhodnocení prostorové autokorelace reziduálních odchylek bylo provedeno pomocí výpočtu *Moranova I* a vizualizace, a to vždy na vrstvě normovaných reziduí pro výsledný regresní model sestavený pro každé z 9 území (viz tab. 16).

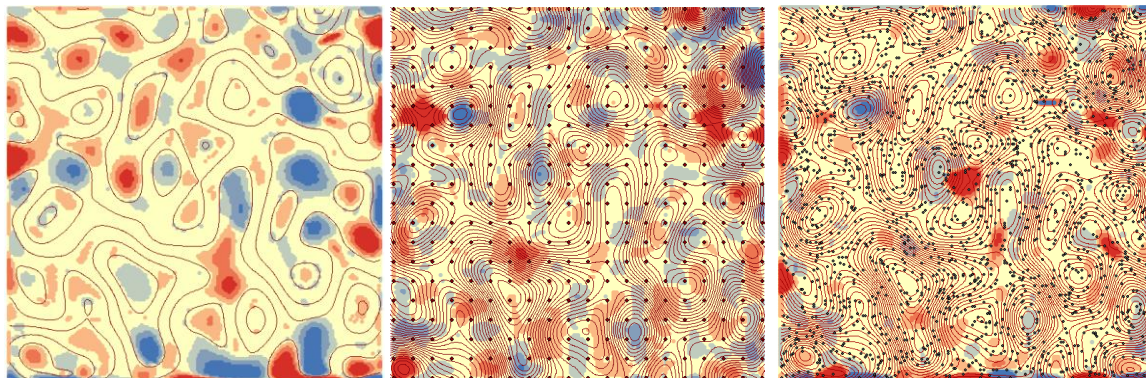
V ArcGIS Desktop, pomocí funkce *Spatial Statistics/Spatial Autocorrelation (Global Moran's I)*, bylo vypočteno Moranovo I, které měří prostorovou autokorelaci na základě vzdálenosti mezi prvky a hodnotami jejich atributů. Výsledkem jsou i odpovídající Z score a p-value, které vyjadřují statistickou významnost Moranova I. Z výsledků prostorové autokorelace (viz tab. 17) vyplývá, že u všech devíti modelů jsou hodnoty Z score a p-value statisticky významné na hladině $\alpha = 0,05$. Kladná hodnota Moranova I okolo hodnoty 1 značí tendenci reziduálních odchylek shlukovat se.

Tab. 17 Výsledné hodnoty Moranova I kritéria, Z score a p-value

Matematický terén 1 x 1 km	Moranovo I	Z score	p-value
<i>Vrstevnice</i>			
1 m	0,801	226,112	0,000
3 m	0,925	260,922	0,000
5 m	0,957	269,942	0,000
<i>Náh. rozmís. body</i>			
5%	0,899	253,923	0,000
10%	0,877	247,608	0,000
55%	0,834	235,676	0,000
<i>Prav. rozmís. body</i>			
15 m	0,933	258,429	0,000
25 m	0,911	245,679	0,000
50 m	0,961	259,120	0,000

Následně byla provedena prostorová analýza reziduí pomocí vizualizace. Vizualizovány byly vrstvy normovaných reziduí pro výsledné regresní modely (dle tab. 16). Grafické podklady pro vizuální analýzu jsou zpracovány v příloze 2, ukázka je na obrázku č. 17. Z výsledků reziduí vyplývá, že chyby nezávisí až tak na sklonu, orientaci a podobně, jak se původně předpokládalo, ale jsou velmi ovlivněny právě rozmístěním a hustotou vstupních dat. Čím je větší vzdálenost mezi vstupními body, tím je markantnější vliv dat. Pokud byly vstupními daty vrstevnice, rezidua tvořily koncentrické shluky, které vrstevnice kopírovaly. Při větším rozestupu vrstevnic dochází

k větším a shlazenějším shlukům, což je způsobeno menším počtem vstupních dat. Vliv vstupních dat na rezidua je viditelný i u pravidelně rozmístěných bodů, nejlépe je to vidět u bodů po 50 m, kdy jsou vstupní data dále od sebe. Rezidua jsou koncentrována spíše do čtverce, než do kruhu jak bylo u vrstevnic. U náhodně rozmístěných bodů nelze pozorovat podobnost shluků, jelikož se jedná o náhodně rozmístěné body, které budou vždy jiné. Platí tedy, že čím větší hustota vstupních dat, tím menší shluky reziduí. U okrajů jednotlivých DEM vznikají větší shluky reziduí, což je způsobeno větší chybou interpolace, pravděpodobně nedostatkem dat.



Obr. 17 Ukázka vizualizace prostorové analýzy reziduí OLS pro matematický model. Vlevo je zobrazeno území DEM vytvořené ze ZIV po 5 m, uprostřed z pravidelně rozmístěných bodů po 50 m, vpravo pak hustota z náhodně rozmístěných bodů 5 %.

Z grafů závislostí reziduí na parametrech z výsledného regresního modelu, kterými jsou vzdálenost, sklon svahu, orientace svahu, orientace svahu reklasifikovaná a celková křivost (viz příloha 3), je dobře patrné, že hodnoty reziduí jsou většinou rozptýleny okolo jedné konstantní hodnoty. Křivku závislosti mezi rezidui a parametry lze kromě křivosti vyjádřit horizontální linií, což znamená, že závislost na parametrech není nebo je velice nízká. To dokazují i výsledky regresních koeficientů β . Výjimkou je celková křivost, jejíž graf nemá horizontální závislost a je volně rozptýlen, proto můžeme říct, že na hodnoty chyb mají největší vliv hodnoty křivosti.

6.2.2 Geograficky vážená regrese - GWR

Celý proces GWR byl prováděn na umělých matematických datech na území o velikosti 1 x 1 km a v prostředí ArcGIS Desktop. Do výpočtu vstupovala jediná bodová vrstva obsahující v atributové tabulce hodnoty chyb jako závislá proměnná a jako nezávislé proměnné hodnoty zkoumaných parametrů z výsledného regresního modelu (celkem 7 parametrů - vzdálenost ke vstupním bodům (u vrstevnic k vertexům), sklon svahu, orientace svahu, orientace svahu reklasifikovaná, celková křivost a také mezi parametry získané pronásobením základních morfometrických parametrů sklonu svahu \times profilové křivosti a orientace svahu \times planární křivosti). Pro výpočet byl použit nástroj *Spatial Statistics Tools/Modeling Spatial Relationships/Geographically Weighted Regression*, kde byl typ prostorového jádra zvolen jako fixní a šířka pásma byla

stanovena na 100 m. Výběr tohoto nastavení je daný tím, že při vstupu tolika vysvětlujících parametrů metoda CV (Cross Validation) a AICc (Akaike Information Criterion) selhává.

Před vlastní analýzou reziduí metody GWR je důležité zhodnotit, zda použití této metody vede k lepším výsledkům než využití metody OLS. Pro srovnání se nabízí využití několika statistik. První je hodnota koeficientu determinace Adjusted R^2 . Druhou statistikou, umožňující kontrolovat vhodnost prostorového modelu, je Akaikeho informační kritérium (Akaike Information Criterion – AICc), které je založeno na odpovídající funkci pravděpodobnosti standardního regresního modelu a na předpokladu vícerozměrné normality. V tomto případě platí, že vyšší explanační sílu modelu GWR před modelem OLS indikuje nižší hodnota AIC (Anselin, 2005). Srovnání vhodnosti modelů umožňuje i pohled na reziduální součet čtverců (sum of square), kdy platí, že s poklesem reziduálního součtu čtverců modelu narůstá vhodnost modelu (Hendl, 2004).

Základní srovnání regresních modelů ukazuje, že model GWR má jednoznačně větší explanační sílu a větší vysvětlenou varianci než model OLS. Zatímco model OLS dokázal vysvětlit velikost hodnot chyb pomocí hodnot zkoumaných parametrů maximálně z 24%, v případě GWR byl nárůst až na 69%. Lepší vlastnosti modelu GWR dokládá i pokles hodnoty AICc a srovnání součtu čtverců reziduí, které v případě modelů GWR vykazují výrazný pokles (viz tab. 18).

Tab. 18 Test zlepšení vhodnosti regresního modelu na umělých matematických datech

Matematický terén 1 x 1 km	OLS		GWR		reziduální součet čtverců	
	Ad. R^2	AICc	Ad. R^2	AICc	OLS - rezidua	GWR - rezidua
<i>Vrstevnice</i>						
1 m	0,237	85786,937	0,595	57398,742	19989,872	10154,482
3 m	0,068	196122,870	0,585	140239,416	315331,375	106361,633
5 m	0,085	258614,633	0,625	191623,795	1504052,011	523692,732
<i>Náh. rozmís. body</i>						
5%	0,067	228983,527	0,525	181190,342	715045,139	347963,700
10%	0,190	217658,018	0,681	161048,664	539621,266	198472,686
55%	0,169	109881,752	0,505	82328,753	36367,806	20143,849
<i>Prav. rozmís. body</i>						
15 m	0,120	172274,651	0,647	95541,310	199253,721	32231,261
25 m	0,111	142410,311	0,689	98003,157	109158,498	36192,680
50 m	0,088	205814,556	0,618	154795,991	601912,901	219592,935

Stejně jako u OLS i zde byla provedena vizualizace prostorové analýzy reziduí jednotlivých modelových území, která je zpracována v příloze 4. Z důvodu zlepšení výsledků regresního modelu dochází k menší četnosti reziduí, ale i zde je vidět vliv vstupních dat. U vrstevnic tvoří rezidua koncentrické shluky, u pravidelně rozmístěných bodů dochází k tomu, že rezidua jsou koncentrována spíše do čtverce, než do kruhu jak bylo u vrstevnic. U náhodně rozmístěných bodů nelze pozorovat podobnost shluků, jelikož se jedná o náhodně rozmístěné body.

7 VERIFIKACE VÝSLEDKŮ NAD DATY Z LETECKÉHO LASEROVÉHO SKENOVÁNÍ

7.1 Korelační analýza

Cílem korelační analýzy byla snaha o nalezení vztahů mezi hodnotami chyb a hodnotami zkoumaných parametrů: nadmořská výška, převýšení, vzdálenost ke vstupním bodům (u vrstevnic k vertexům), sklon svahu, orientace svahu, orientace svahu reklasifikovaná, planární, profilová a celková křivost, extremita minima, maxima a průměru a také parametry získané pronásobením sklonu svahu \times profilové křivosti a orientace svahu \times planární křivosti. Pro výpočet byl použit neparametrický Spearmanův pořadový korelační koeficient, který je podrobněji popsán v kapitole 3.4. Celý proces byl prováděn na datech z leteckého laserového skenování na území o velikosti 1,1 x 1,4 km a v prostředí statistického programu R.

Při hodnocení těsnosti vztahů mezi hodnotami chyb a hodnotami zkoumaných parametrů (viz tab. 19) byla kromě devíti případů prokázána existence statisticky významného vztahu. U většiny případů, dle korelačního koeficientu, je těsnost těchto vztahů nízká nebo mírná ($r_s < 0,5$). Z toho vyplývá, že hodnoty zkoumaných parametrů neovlivňují velikost hodnot chyb. Díky prokázané, sice nízké či mírné, závislosti mezi zkoumanými proměnnými je však možné vypočítat regresní analýzu pomocí vícenásobného regresního modelu.

7.2 Regresní analýza

Základním cílem regresní analýzy je nalezení regresní funkce, která co nejlépe vystihuje zkoumané vztahy a vyjádření této funkce matematicky. K určení vztahu mezi hodnotami chyb a hodnotami či změnami hodnot zkoumaných parametrů odvozenými z DEM byl použit lineární model OLS regrese a poté model geograficky vážené regrese.

7.2.1 Vícenásobná lineární OLS regrese – metoda nejmenších čtverců

Celý proces výpočtu OLS regrese byl prováděn na datech z leteckého laserového skenování na území o velikosti 1,1 x 1,4 km a v prostředí ArcGIS Desktop. Do výpočtu vstupovala jediná bodová vrstva obsahující v atributové tabulce hodnoty chyb jako závislá proměnná a jako nezávislé proměnné hodnoty zkoumaných parametrů (celkem 14 parametrů - nadmořská výška, převýšení, vzdálenost ke vstupním bodům, sklon svahu, orientace svahu, orientace svahu reklasifikovaná, planární, profilová a celková křivost, extremita minima, maxima a průměru a také parametry získané pronásobením sklonu svahu \times profilové křivosti a orientace svahu \times planární křivosti). V první variantě regresního modelu jsou hodnoty koeficientu determinace Adjusted R^2 velmi nízké (viz tab. 20). Stejně jako u matematického modelu i zde dochází k multikolinearitě mezi parametry. Proces výpočtu každého regresního modelu se tedy několikrát opakoval dle principu postupné zpětné regrese, důvodem byla snaha o zlepšení regresního modelu.

Tab. 19 Hodnoty Spearmanova korelačního koeficientu mezi hodnotami chyb a hodnotami zkoumaných parametrů

Laser-scan 1,1 x 1,4 km	Zkoumané parametry														
		nadmořská výška	převýšení	vzdálenost	sklon	orientace	orientace reklasifikovaná	křivost celková	křivost planární	křivost profilová	sklon x profilová křivost	orientace x planární křivost	extremita maxima	extremita minima	extremita průměru
<i>Vrstevnice</i>															
1 m	chyby	0,024	0,005	0,013	0,016	0,099	0,093	0,313	0,221	-0,282	-0,296	0,203	-0,018	-0,020	-0,022
3 m	chyby	-0,003	0,012	0,018	0,022	0,002	-0,007	0,120	0,106	-0,070	-0,091	0,094	0,010	0,010	0,008
5 m	chyby	-0,009	-0,038	0,091	-0,031	0,033	0,019	0,071	0,140	0,040	0,023	0,135	0,011	0,000	0,004
<i>Náhodně rozmístěné body</i>															
5%	chyby	0,047	0,007	0,044	0,009	0,191	0,212	0,183	0,141	-0,121	-0,129	0,139	-0,044	-0,047	-0,047
10%	chyby	0,014	0,007	0,016	0,036	0,108	0,115	0,212	0,158	-0,168	-0,182	0,148	-0,011	-0,013	-0,013
55%	chyby	-0,013	0,017	0,005	0,025	0,049	0,053	0,411	0,321	-0,361	-0,370	0,303	0,016	0,013	0,014
<i>Pravidelně rozmístěné body</i>															
15 m	chyby	-0,007	0,007	-0,002	0,006	0,050	0,056	0,274	0,207	-0,230	-0,236	0,188	0,010	0,007	0,008
25 m	chyby	0,008	0,016	-0,001	0,004	0,173	0,186	0,197	0,158	-0,142	-0,154	0,148	-0,028	-0,009	-0,008
50 m	chyby	0,002	0,067	0,010	0,068	0,075	0,078	0,168	0,161	-0,056	-0,063	0,139	0,009	0,023	0,015

Těsnost závislosti: nulová nízká mírná

Tab. 20 Výsledné hodnoty prostého členu α , regresních koeficientů a koeficientu determinace Adjusted R^2 získané pomocí vícenásobné lineární OLS regrese

Laser-scan 1,1 x 1,4 km	α	Regresní koeficienty nezávisle proměnných														Ad. R^2
		β_1 nadmořská výška	β_2 převýšení	β_3 vzdálenost	β_4 sklon	β_5 orientace	β_6 orientace reklasifikovaná	β_7 křivost celková	β_8 křivost planární	β_9 křivost profilová	β_{10} sklon x profilová křivost	β_{11} orientace x planární křivost	β_{12} extremita maxima	β_{13} extremita minima	β_{14} extremita průměru	
<i>Vrstevnice</i>																
1 m	-0,021	-0,095	-0,022	0,001	0,015	0,000	0,007	-0,013	0,225	0,011	-0,035	0,000	-0,104	0,066	-0,057	0,318
3 m	0,260	-0,032	0,014	0,004	0,016	0,000	-0,004	-0,242	0,376	-0,091	-0,035	0,000	-0,013	0,012	-0,030	0,120
5 m	0,380	-0,004	0,092	0,009	0,015	0,000	-0,016	-0,343	0,593	0,102	-0,047	0,000	0,057	-0,015	-0,046	0,097
<i>Náh. rozmís. body</i>																
5%	-0,043	0,149	0,291	-0,002	-0,051	0,000	0,051	-0,478	0,819	-0,125	-0,051	0,000	0,142	-0,189	0,196	0,085
10%	-0,073	-0,053	0,030	0,001	0,019	0,000	0,020	-0,103	0,381	0,034	-0,047	-0,001	-0,051	0,007	-0,009	0,098
55%	-0,026	-0,120	0,016	0,001	0,009	0,000	0,010	0,081	0,140	0,203	-0,044	0,000	-0,142	0,071	-0,048	0,293
<i>Prav. rozmís. body</i>																
15 m	0,100	-0,047	0,085	0,002	0,012	0,000	0,010	-0,017	0,209	0,078	-0,041	0,000	-0,047	-0,019	0,020	0,129
25 m	0,189	0,029	0,088	0,001	0,014	0,000	0,032	-0,189	0,492	0,158	-0,063	-0,001	0,041	-0,089	0,078	0,106
50 m	0,183	-0,020	0,285	-0,001	-0,028	0,001	0,020	-0,420	1,417	-0,188	-0,023	-0,002	0,038	0,037	-0,093	0,043

vysoký faktor VIF (Variance Inflation Factor) VIF > 7,5

Prvním krokem bylo odstranění parametrů na základě vysokého faktoru VIF (Variance Inflation Factor). Pokud byla některá z vysvětlujících proměnných ohodnocena $VIF > 7,5$, byla z modelu odstraněna. Limitní hodnota je dle doporučení ArcGIS Resource Center.

Dalším krokem pro zlepšení regresního modelu bylo odstranění parametrů dle výsledků vzájemných korelací mezi proměnnými (viz tab. 9). Tam, kde se vyskytla význačná až velmi vysoká korelace, vstoupil do následného regresního modelu vždy jen jeden parametr s vyšší hodnotou korelace s hodnotami chyb (dle tab. 19) z důvodu předpovídání větší chyby.

Výsledné tabulky OLS regrese a jejich postupů u dat z leteckého laserového skenování jsou umístěny v příloze 5.

Dalším a posledním přístupem, který vedl ke zlepšení a zároveň k výslednému regresnímu modelu metodou OLS, bylo odstranění kombinací parametrů dle vysokého faktoru VIF a dle výsledků vzájemných korelací mezi proměnnými. Za výsledný regresní model, kde bylo dosaženo nejlepších výsledků koeficientů determinace, lze považovat model s parametry uvedenými v tabulce 21. Jako nezávisle proměnné vstupovaly parametry vzdálenosti ke vstupním bodům (u vrstevnic k vertexům), sklonu svahu, orientace svahu, orientace svahu reklasifikované, celkové křivosti a parametry získané pronásobením základních morfometrických parametrů sklon svahu x profilová křivost a orientace svahu x planární křivost (celkem 7 parametrů). U výsledků hodnot koeficientů determinace R^2 , oproti předchozímu odstranění parametrů, dochází k mírnému zlepšení koeficientů determinace maximálně do 0,27, což znamená, že změny hodnot chyb nelze pomocí hodnot zkoumaných parametrů vysvětlit z více jak 27%. Z celkového pohledu na regresní koeficienty β vyplývá, že výrazně vyšší hodnoty jsou u křivosti, které mají vyšší efekt na hodnoty chyb.

Tab. 21 Výsledné hodnoty prostého členu α , regresních koeficientů a koeficientu determinace Adjusted R^2 po odstranění kombinací parametrů dle vysokého faktoru VIF a dle výsledků vzájemných korelací mezi proměnnými získané pomocí vícenásobné lineární OLS regrese → výsledný regresní model

Laser-scan 1 x 1 km	α								Ad. R^2
		β_1 vzdálenost	β_2 sklon	β_3 orientace	β_4 orientace reklasifikovaná	β_5 křivost celková	β_6 sklon x profilová křivost	β_7 orientace x planární křivost	
<i>Vrstevnice</i>									
1 m	-0,051	0,001	0,001	0,000	0,006	0,020	-0,024	0,000	0,269
3 m	-0,036	0,005	0,003	0,000	-0,005	-0,048	-0,025	0,001	0,100
5 m	-0,092	0,010	0,004	0,000	-0,014	-0,145	-0,027	0,002	0,070
<i>Náh. rozmís. body</i>									
5%	-0,303	-0,001	0,006	0,000	0,063	0,023	-0,025	0,001	0,061
10%	-0,155	0,002	0,005	0,000	0,021	0,013	-0,026	0,000	0,075
55%	-0,046	0,001	0,001	0,000	0,010	-0,041	-0,026	0,001	0,195
<i>Prav. rozmís. body</i>									
15 m	-0,077	0,002	0,004	0,000	0,012	-0,009	-0,023	0,000	0,085
25 m	-0,233	0,002	0,005	0,000	0,035	-0,078	-0,036	0,001	0,071
50 m	-0,319	-0,001	0,018	0,000	0,028	0,179	0,004	0,001	0,026

Nutnou součástí hodnocení výsledků regrese je také výpočet testu celkové statistické významnosti (*Joint Wald Statistic*), stacionarity (*Koenker Statistic*) a zkreslení modelu (*Jarque-Bera Statistic*) a zhodnocení prostorové autokorelace reziduálních odchylek (pomocí shlukování). Z výsledků těchto testů bylo zjištěno, že všechny vytvořené modely lze považovat za statisticky významné, avšak nestacionární a zkreslené (ve všech případech byla $p\text{-value} < 0,05$), stejně jako tomu bylo u matematického modelu.

Zhodnocení prostorové autokorelace reziduálních odchylek bylo provedeno pomocí *Moranova I* a to na vrstvě normovaných reziduí. Z výsledků vyplývá, že u všech modelů pro území s různou hustotou a rozmístěním vstupních bodů jsou hodnoty *Z score* a *p-value* statisticky významné na hladině $\alpha = 0,05$. Tudíž kladná hodnota *Moranova I* okolo hodnoty 1 značí tendenci ke shlukování.

Následně byla provedena prostorová analýza reziduí na vrstvě normovaných reziduí pro každý výsledný regresní model, do kterého vstupovalo 7 parametrů - vzdálenost ke vstupním bodům (u vrstevnic k vertexům), sklon svahu, orientace svahu, orientace svahu reklasifikovaná, celková křivost a parametry získané pronásobením základních morfometrických parametrů sklon \times profilová křivost a orientace \times planární křivost. Vizualizace prostorové analýzy reziduí jednotlivých modelových území je zpracována v příloze 6. Výskyt reziduí nad daty z leteckého laserového skenování také kopíruje hustotu a rozmístění vstupních dat, ale není to tak výrazné jako u matematického modelu, kde byla jiná hustota vstupních dat. U vrstevnic rezidua tvoří koncentrické shluky, ale jelikož data z laser scanu měla hodně husté vstupní bodové pole (viz tab. 22), vrstevnice jsou kostrbatější a tím pádem shluky neuhlazené. Pravidelně rozmístěné body mají rezidua koncentrována spíše do čtverce než do kruhu jak bylo u vrstevnic. U náhodně rozmístěných bodů nelze pozorovat podobnost shluků, jelikož se jedná o náhodně rozmístěné body, které budou vždy jiné. Platí tedy, že čím větší hustota vstupních dat, tím menší shluky reziduí.

Tab. 22 Hustota počtu vstupních bodů/m² u laser scanu

	Hustota (počet vstupních bodů/m ²)
Prav. rozmís. body	
15 m	0,007
25 m	0,003
50 m	0,001
Náh. rozmís. body	
5%	0,003
10%	0,006
55%	0,035
Vrstevnice	
1 m	0,146
3 m	0,048
5 m	0,029

7.2.2 Geograficky vážená regrese - GWR

Celý proces zpracování GWR byl prováděn na reálných datech z laser scanu na území o velikosti 1,1 x 1,4 km, kde do výpočtu vstupovala jediná bodová vrstva obsahující v atributové tabulce hodnoty chyb jako závislá proměnná a jako nezávislé proměnné hodnoty zkoumaných parametrů z výsledného regresního modelu (celkem 7 parametrů - vzdálenost ke vstupním bodům (u vrstevnic k vertexům), sklon svahu, orientace svahu, orientace svahu reklasifikovaná, celková křivost a také parametry získané pronásobením základních morfometrických parametrů sklonu svahu × profilové křivosti a orientace svahu × planární křivosti).

Před vlastní analýzou reziduí metody GWR je důležité zhodnotit, zda použití této metody vede k lepším výsledkům než využití metody OLS. Základní srovnání regresních modelů ukazuje, že model GWR má jednoznačně větší explanační sílu a větší vysvětlenou varianci než model OLS. Zatímco model OLS dokázal vysvětlit velikost hodnot chyb pomocí hodnot zkoumaných parametrů maximálně z 27%, v případě GWR byl nárůst až na 47%. Lepší vlastnosti modelu GWR dokládá i pokles hodnoty AICc a srovnání součtu čtverců reziduí, které v případě modelů GWR vykazují výrazný pokles (viz tab. 23).

Tab. 23 Test zlepšení vhodnosti regresního modelu na reálných datech z laser scanu

Laser-scan 1,1 x 1,4 km	OLS		GWR		reziduální součet čtverců	
	Ad. R ²	AICc	Ad. R ²	AICc	OLS - rezidua	GWR - rezidua
<i>Vrstevnice</i>						
1 m	0,269	281595,999	0,372	83340,965	320333,365	83407,762
3 m	0,100	331066,041	0,243	65251,201	701612,639	113576,470
5 m	0,070	382700,753	0,434	111073,175	1590346,320	390297,399
<i>Náh. rozmís. body</i>						
5%	0,061	408736,572	0,470	99927,111	2402661,083	366820,212
10%	0,075	352078,852	0,320	101531,587	978161,176	212676,418
55%	0,195	288834,858	0,396	101434,919	359070,789	91570,227
<i>Prav. rozmís. body</i>						
15 m	0,085	313770,062	0,260	100788,567	581048,712	171124,875
25 m	0,071	351994,415	0,340	86587,415	1147264,895	215189,444
50 m	0,026	404842,888	0,349	82738,877	3976090,689	506219,640

Stejně jako u OLS i zde byla provedena vizualizace prostorové analýzy reziduí jednotlivých modelových území, která je zpracována v příloze 7. Z důvodu zlepšení výsledků regresního modelu dochází k menší četnosti reziduí, ale i zde je vidět částečně vliv vstupních dat.

8 VÝSLEDKY

Hlavním výsledkem diplomové práce je nalezení hlavních parametrů, které lze odvodit z DEM a z povahy vstupních dat (např. rozmístění či hustota dat) a lze je použít pro odhad chyb v DEM. Dalším, neméně důležitým, výsledkem je sestavení regresních modelů umožňujících odhad chyby v DEM na základě těchto vybraných parametrů. Nedílnou součástí diplomové práce je textová část a její přílohy a www stránky.

V praktické části bylo nejprve provedeno hodnocení kvality výškových modelů jak pomocí globálních neprostorových, tak lokálních prostorových charakteristik, kde bylo využito vážené pořadí a lokální shlukové analýzy LISA. Dále bylo testováno, zda existuje závislost mezi velikostí chyb v DEM a hodnotami zkoumaných parametrů, k tomu byla využita korelační a regresní analýza. Výsledky jsou založené na uměle generovaných datech a na reálných datech z leteckého laserového skenování části obce Vysoké Pole.

Z analýzy získaných výsledků hodnocení kvality DEM pomocí váženého pořadí vyplývá, že nejvhodnějšími vstupními daty se jeví husté rozmístění vrstevnic se ZIV po 1 m, které dosáhlo nejlepšího hodnocení. Naopak rozmístění vrstevnic se ZIV po 5 m, kde s klesající hustotou vstupních dat dochází i ke vzniku větších rozdílů mezi jednotlivými DEM, se jeví jako nejméně vhodné. Vážené pořadí je značně ovlivněno vysokou hustotou vstupních bodů v závislosti na velikosti území a nelze stanovit ostrou hranici mezi kvalitními a nekvalitními DEM, ale lze vybrat nejkvalitnější DEM dle použitých ukazatelů pro další zpracování. Výběru vhodného DEM by tedy vždy mělo předcházet vytvoření více DEM a sestavení váženého pořadí. Z prostorového hodnocení pomocí LISA vychází, že v DEM označovaném jako kvalitnější se nachází menší množství statisticky významných hodnot chyb než v DEM označených jako méně kvalitní. Toto hodnocení nelze nikdy použít samostatně, jelikož by mohlo dojít k milné interpretaci výsledků, proto je nutné zaměřit se i na vybrané globální hodnoty ukazatelů (např. rozsah hodnot chyb či průměrná hodnota chyb).

Druhá část je věnována hodnocení závislosti mezi hodnotami chyb v DEM a hodnotami zkoumaných parametrů (nadmožská výška, převýšení, vzdálenost ke vstupním bodům (u vrstevnic k vertexům), sklon svahu, orientace svahu, orientace svahu reklasifikovaná, planární, profilová a celková křivost, extremita minima, maxima a průměru a také parametry získané pronásobením základních morfometrických parametrů sklonu svahu \times profilové křivosti a orientace svahu \times planární křivosti), kde pomocí Spearmanova korelačního koeficientu byla nalezena nízká nebo mírná závislost ($r_s < 0,5$). Význačná závislost ($r_s \in < 0,5; 0,7$) byla prokázána jen v několika málo případech mezi hodnotami chyb a hodnotami celkové křivosti, planární a profilové křivosti. Jelikož při analýze korelace není stanoveno, která z proměnných je vysvětlovaná a která vysvětlující, lze konstatovat, že hodnoty křivostí lze alespoň částečně považovat za indikátory výskytu chyb. Díky prokázané závislosti mezi zkoumanými proměnnými bylo možné vypočítat regresní analýzu.

Z výsledků vícenásobné lineární regrese vyplývá, že velikost hodnot chyb nelze u matematického DEM a ani u laser scanu dostatečně vysvětlit pomocí zkoumaných parametrů. Napovídají tomu jak velmi nízké hodnoty koeficientů determinace (většinou $R^2 < 0,2$), tak výsledky testů celkové statistické významnosti (*Joint Wald Statistic*), stacionarity (*Koenker Statistic*), zkreslení modelů (*Jarque-Bera Statistic*) a výsledky zhodnocení prostorové autokorelace reziduálních odchylek. Ale i přes velmi nízké hodnoty koeficientů determinace byl sestaven výsledný regresní model, kde bylo dosaženo nejlepších, avšak nízkých, výsledků těchto koeficientů. Jako nezávisle proměnné vstupovaly parametry vzdálenosti ke vstupním bodům (u vrstevnic k vertexům), sklonu svahu, orientace svahu, orientace svahu reklasifikované, celkové křivosti a parametry získané pronásobením základních morfometrických parametrů sklon svahu x profilová křivost a orientace svahu x planární křivost. Z výsledků těchto testů vyplývá, že všechny vytvořené modely lze považovat za statisticky významné, avšak nestacionární a zkreslené (ve všech případech byla p-value < 0,05). Z důvodu nestacionarity modelu bylo tedy vhodnější použít prostorový typ regrese – geograficky váženou regresi (GWR).

Základní srovnání regresních modelů ukazuje, že model GWR jak u matematického modelu, tak u laser scanu má jednoznačně větší explanační sílu a větší vysvětlenou varianci než model OLS. Zatímco model OLS dokázal vysvětlit velikost hodnot chyb pomocí zkoumaných parametrů maximálně z 30%, v případě GWR byl nárůst až na 70%. Lepší vlastnosti modelu GWR dokládá i pokles hodnoty AICc a srovnání součtu čtverců reziduí, které v případě modelů GWR vykazují výrazný pokles.

Textová část obsahuje rešerše zaměřené na podobnou problematiku, jako je zadané téma diplomové práce. Jsou zde popsány faktory ovlivňující vznik chyb v DEM, způsoby hodnocení kvality DEM a analýza vztahů mezi parametry.

O diplomové práci byly vytvořené www stránky dostupné ze stránek Katedry geoinformatiky. Obsahem těchto stránek je stručná charakteristika diplomové práce (cíle, postup, metody, výsledky).

9 DISKUZE

Cílem diplomové práce je nalezení hlavních parametrů, které lze odvodit z DEM a z povahy vstupních dat (např. rozmístění či hustota dat) a lze je použít pro odhad chyb v DEM. Dalším, neméně důležitým, cílem je sestavení regresních modelů umožňujících odhad chyby v DEM na základě těchto vybraných parametrů.

Na samém začátku této práce bylo nejdůležitějším úkolem zvolit vhodnou interpolační metodu pro tvorbu DEM. Každá interpolační metoda má své klady a zápory a každé území má svá charakteristická specifika, proto vybrat plošně nejvhodnější metodu interpolace je obtížné a časově náročné. Z tohoto důvodu bylo v diplomové práci počítáno převážně jen s jednou interpolační metodou a to s nástrojem implementovaným v ArcGIS - Topo To Raster, který kombinuje několik interpolačních metod. Byla testována i jiná, běžně dostupná, interpolační metoda a to běžný kriging - semivariogram - sférický, ale u výsledků nedošlo ke zlepšení intenzity vztahů mezi hodnotami chyb a hodnotami zkoumaných parametrů. Proto bylo v diplomové práci od testování na dalších interpolačních metodách upuštěno.

První částí práce bylo hodnocení kvality DEM a z něj odvozených charakteristik. Neprostorové hodnocení metrické přesnosti bylo založeno na výpočtu váženého pořadí dle vybraných neprostorových ukazatelů - střední kvadratické chyby, celkové absolutní chyby, minimální a maximální chyby, průměru a směrodatné odchylky. Dle autora práce tyto charakteristiky umožňují dostatečně popsat přesnost DEM. Bylo by ale jistě možné použít i jiné neprostorové ukazatele např. hammock index, průměrnou absolutní chybu nebo rozptyl absolutní chyby. K prostorovému hodnocení kvality DEM bylo využito lokální shlukové analýzy LISA, která vyžaduje na vstupu pouze kladné hodnoty. Z toho důvodu byly hodnoty chyb získané odečtením vytvořených DEM od referenčního převedeny na absolutní hodnoty.

Druhá část práce se zabývala zhodnocením vztahů mezi hodnotami chyb v DEM a hodnotami zkoumaných parametrů pomocí korelační a regresní analýzy. Z nízkých hodnot korelačních koeficientů, koeficientů determinace regresních modelů i z testů celkové statistické významnosti, stacionarity, zkruslení výsledných modelů a z výsledků zhodnocení prostorové autokorelace reziduálních odchylek je patrné, že tyto parametry nejsou dostačující. Ale i přes velmi nízké hodnoty koeficientů determinace byl naplněn cíl práce a sestaven výsledný regresní model, kde bylo dosaženo nejlepších, avšak nízkých, výsledků těchto koeficientů. V následujícím výzkumu by tedy bylo dobré se zaměřit na nalezení dalších vysvětlujících proměnných, které by vedly k lepšímu vysvětlení změn hodnot chyb.

Uplatnění výsledků diplomové práce lze hledat zejména u uživatelů, kteří ve své práci využívají DEM a z něj odvozené parametry. Jedná se tedy o geomorfology, územní plánovače, krajinné ekology a další. Výsledky práce by se jim měly stát návodem, jak vytvořit kvalitní DEM z hlediska metrické přesnosti a také jak nalézt místa (shluky) se statisticky významnými hodnotami vysokých chyb.

10 ZÁVĚR

Úkolem této diplomové práce bylo nalézt hlavní parametry, které lze odvodit z DEM a z povahy vstupních dat (např. rozmístění či hustota dat) a lze je použít pro odhad chyb v DEM. Dalším, neméně důležitým, výsledkem bylo sestavení regresních modelů umožňujících odhad chyby v DEM na základě těchto vybraných parametrů. Pro tento účel bylo testováno na uměle generovaných datech a na datech z leteckého laserového skenování části obce Vysoké Pole.

Základní linie zpracování diplomové práce postupovala od studia literatury, kde jsou popsány faktory ovlivňující vznik chyb v DEM, způsoby hodnocení kvality DEM a analýza vztahů mezi parametry. Dále přes tvorbu DEM modelových území a jejich hodnocení jak pomocí globálních neprostorových, tak lokálních prostorových charakteristik, až ke snaze nalézt a matematicky definovat vztahy mezi hodnotami chyb v DEM a hodnotami zkoumaných parametrů odvozenými z DEM. K tomu byla využita korelační a regresní analýza.

Na samém začátku této práce bylo třeba zpracovat data a vytvořit tak přesný digitální výškový model. Pro tvorbu DEM byla použita interpolační metoda Topo to Raster, která kombinuje několik interpolačních metod, aby vytvořila hydrologicky korektní model terénu. Pro každou datovou sadu bylo vygenerováno 9 gridů (celkem 18 gridů) s různou hustotou a rozmístěním vstupních bodů. Celkem bylo vytvořeno 252 gridů. Následné hodnocení kvality vytvořených DEM bylo založeno na výpočtu váženého pořadí dle vybraných neprostorových ukazatelů - střední kvadratické chyby (RMSE), celkové absolutní chyby (AE), minimální a maximální chyby, průměru a směrodatné odchylky (SD). K prostorovému hodnocení kvality DEM bylo využito lokální shlukové analýzy LISA, kde výsledkem je vymezení shluků statisticky významných chyb a pomocí níž lze sledovat prostorové rozmístění chyb. Existence a těsnost vztahů mezi hodnotami chyb a hodnotami zkoumaných parametrů byla ověřována pomocí Spearmanova pořadového korelačního koeficientu. Vícenásobná lineární OLS regrese tak i geograficky vážená regrese (GWR) byla poté využita ke zhodnocení a matematickému vyjádření vztahů mezi hodnotami chyb jako závisle proměnnou a hodnotami zkoumaných parametrů (nadmořská výška, převýšení, vzdálenost ke vstupním bodům (u vrstevnic k vertexům), sklon svahu, orientace svahu, orientace svahu reklasifikovaná, planární, profilová a celková křivost, extrimita minima, maxima a průměru a také parametry získané pronásobením základních morfometrických parametrů sklonu svahu \times profilové křivosti a orientace svahu \times planární křivosti) odvozenými z DEM jako nezávisle proměnnými. Zpracování diplomové práce proběhlo především v prostředí ArcGIS Desktop, částečně také ve statistickém softwaru R.

Jedním z hlavních výsledků práce je hodnocení kvality DEM pomocí váženého pořadí, kde se nejvhodnějšími vstupními daty jeví husté rozmístění vrstevnic se ZIV po 1 m, které dosáhlo nejlepšího hodnocení. Naopak rozmístění vrstevnic se ZIV po 5 m se jeví jako nejméně vhodné. Výběru vhodného DEM by tedy vždy mělo předcházet

vytvoření více DEM a sestavení váženého pořadí. Dalším výsledkem práce je hodnocení kvality DEM pomocí lokální shlukové analýzy LISA, kde v DEM označovaném jako kvalitnější se nachází menší množství statisticky významných hodnot chyb než v DEM označených jako méně kvalitní. Toto hodnocení nelze nikdy použít samostatně, jelikož by mohlo dojít k mylné interpretaci výsledků.

Významným výsledkem práce je i zhodnocení vztahů mezi hodnotami chyb v DEM a hodnotami zkoumaných parametrů pomocí korelační a regresní analýzy. Korelační analýza ve většině případů potvrdila existenci alespoň mírné závislosti mezi hodnotami chyb a hodnotami zkoumaných parametrů. Význačná závislost byla prokázána jen v několika málo případech mezi hodnotami chyb a hodnotami celkové křivosti, planární a profilové křivosti. Lze konstatovat, že hodnoty křivosti lze alespoň částečně považovat za indikátory výskytu chyb. Následně z výsledků vícenásobné lineární regrese vyplývá, že velikost hodnot chyb nelze u matematického DEM a ani u laser scanu dostatečně vysvětlit pomocí zkoumaných parametrů. Napovídají tomu jak velmi nízké hodnoty koeficientů determinace, tak výsledky testů celkové statistické významnosti, stacionarity, zkreslení modelů a výsledky zhodnocení prostorové autokorelace reziduálních odchylek. Ale i přes velmi nízké hodnoty koeficientů determinace byl sestaven výsledný regresní model, kde bylo dosaženo nejlepších, avšak nízkých, výsledků těchto koeficientů. Jako nezávisle proměnné vstupovaly parametry vzdálenosti ke vstupním bodům (u vrstevnic k vertexům), sklonu svahu, orientace svahu, orientace svahu reklasifikované, celkové křivosti a parametry získané pronásobením základních morfometrických parametrů sklon svahu x profilová křivost a orientace svahu x planární křivost.

Posledním z hlavních výsledků práce je srovnání regresních modelů, které ukazují že model GWR jak u matematického modelu, tak u laser scanu má jednoznačně větší explanační sílu a větší vysvětlenou varianci než model OLS. Zatímco model OLS dokázal vysvětlit velikost hodnot chyb pomocí zkoumaných parametrů maximálně z 30%, v případě GWR byl nárůst až na 70%. Lepší vlastnosti modelu GWR dokládá i srovnání součtu čtverců reziduí, které v případě modelů GWR vykazují výrazný pokles.

Uplatnění výsledků diplomové práce lze hledat zejména u uživatelů, kteří ve své práci využívají DEM a z něj odvozené parametry. Jedná se tedy o geomorfology, územní plánovače, krajinné ekology a další. Výsledky práce by se jim měly stát návodem, jak vytvořit kvalitní DEM z hlediska metrické přesnosti a také jak nalézt místa (shluky) se statisticky významnými hodnotami vysokých chyb.

O diplomové práci byly vytvořeny webové stránky, které jsou umístěny na serveru Katedry geoinformatiky. Celá práce byla rovněž v elektronické podobě umístěna na DVD, které je k práci přiloženo.

POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE

ANDĚL, J. *Statistické metody*. 1. vydání. Praha: MATFYZPRESS, 1993. 246 s.

ANSELIN, L. *Exploring Spatial Data with GeoDaTM : A Workbook*. Spatial Analysis Laboratory. Department of Geography. Urbana-Champaign : University of Illinois, 2005, s. 244. Dostupné z: <<https://geodacenter.asu.edu/system/files/geodaworkbook.pdf>>.

ANSELIN, L. *Local indicators of spatial association – LISA*. *Geographical Analysis*, 1995, roč. 27, s. 93-115.

ArcGIS Resource Center. *Web-based help* [online]. [cit. 2015-03-06]. Dostupné z: <<http://resources.arcgis.com/en/help/main/10.1/index.html#/00qn0000013t000000>>.

BAYER, T. *Digitální modely terénu*. [Web.natur.cuni.cz](http://web.natur.cuni.cz). [online] 2008 [cit. 2014-12-08]. Dostupné z: <<http://web.natur.cuni.cz/~bayertom/IM/idm5.pdf>>.

CARLISLE, B. H. *Modelling the Spatial Distribution of DEM Error*. *Transaction in GIS*, 2005, roč. 9, č. 4, s. 521-540

CARLISLE, B. H. *Digital elevation model quality and uncertainty in DEM-based spatial modelling*. PhD dissertation. University of Greenwich, UK, 2002. [cit. 2014-12-08] Dostupné z: <http://www.numyspace.co.uk/~unn_szbc1/PhD/index.htm>.

CARLISLE, B. H. *The highs and lows of Digital Elevation Model (DEM) error – developing a spatially distributed DEM error model*. In *Proceedings of the 5th International Conference on GeoComputation*. Greenwich, United Kingdom: University of Greenwich, 2000.

DRÁPELA K. *Využití geograficky vážené regrese v lesnických modelech*. In *Sborník Biometrické metody a modely v současné vědě a výzkumu*. Brno: ÚKZÚZ, 2006. s. 81-90. ISBN 80-86548-89-9.

ERDOGAN S. *Modelling the spatial distribution of DEM error with geographically weighted regression: An experimental study*. *Computer & Geosciences*, 2010, roč. 36, s. 34-43.

ERDOGAN S. *A comparison of interpolation methods for producing digital elevation models at the field scale*. *Earth surface processes and landforms*, 2009, roč. 34, s. 366-376.

EVANS, I. S. *General geomorphometry, derivatives of altitude, and descriptive statistics*. In: *Spatial Analysis in Geomorphology*. Londýn: Methuen, 1972. s. 17-90.

FENCÍK, R., VAJSÁBLOVÁ, M. *Parameters of interpolation methods of creation of digital model of landscape*. In Proceeding of the 9th AGILE Conference on Geographic Information Science. Visegrád, Hungary, 2006. s. 374-381.

FISHER, P. F., TATE, N. J. *Causes and consequences of error in digital elevation models*. Progress in Physical Geography, 2006, s. 467–489, ISSN 03091333.

HENDL, J. *Přehled statistických metod zpracování dat. Analýza a metaanalýza dat*. Praha:Portál, 2004, 583 s. ISBN 80-7178-820-1.

JENNESS, J. *DEM Surface Tools for ArcGIS*. Jenness Enterprises. [online]. 2013, [cit. 2015-03-25]. Dostupné z: <http://www.jennessent.com/arcgis/surface_area.htm>.

KRCHO, J. *Morfometrická analýza a digitálne modely georeliéfu*. Slovenská akadémia vied, Bratislava, 1990, 432 s.

LI, J., HEAP, A.D. *A Review of Spatial Interpolation Methods for Environmental Scientists*. Geoscience Australia, 2008, 137 s. ISBN 978-1-921498-28-2.

LI, Z.; ZHU, Q.; Gold, C. *Digital Terrain Modeling: Principles and Methodology*. CRC Press, 2005. 318 s. ISBN 0-415-32462-9.

Lidar Analysis in ArcGIS 9.3.1 for Forestry Applications. [online]. ESRI, June 2010 [cit. 2014-12-09]. Dostupné z: <<http://www.esri.com/library/whitepapers/pdfs/lidar-analysis-forestry.pdf>>.

MOORE, I. D., GRAYSON, R. B., LANDSON, A. R. *Digital Terrain Modelling: A Review of Hydrological, Geomorphological, and Biological Applications*. Hydrological Processes, 1991, roč. 5, s. 3-30.

OKSANEN, J. *Digital elevation model error in terrain analysis*. Helsinki: Faculty of Science, University of Helsinki, Helsinki University Press, Academic Dissertation in Geography, 2006. 51 s. ISBN 978-952-10-3350-6.

OKSANEN, J., SARJAKOSKI, T. *Error propagation of DEM-based surface derivatives*. Computers & Geosciences, 2005, roč. 31, s. 1015-1027.

RAPANT, Petr. *Geoinformatika a geoinformační technologie*. 1. vyd. Ostrava: VŠB - Technická univerzita Ostrava, 2006, 463 s. ISBN 80-248-1264-9.

SPURNÁ, P. *Prostorová autokorelace – všudypřítomný jev při analýze prostorových dat?* Sociologický časopis, 2008, roč. 44, č. 4, s. 767-787.

SVOBODOVÁ, J. *Hodnocení kvality digitálních výškových modelů pro environmentální aplikace*. Disertační práce, Katedra fyzické geografie a geoekologie, Přírodovědecká fakulta, Ostravska Univerzita v Ostravě, 2011, 142 s.

SVOBODOVÁ, Jana: *Kartografické metody vizualizace DEM pro hodnocení jeho kvality*. Geodetický a kartografický obzor, 2011, ročník 57/99, č. 10, 249-254 s.

SVOBODOVÁ, J. *Hodnocení přesnosti digitálních modelů reliéfu*. [online]. Geomorphologia Slovaca et Bohemica 1/2008, Bratislava, 6 s. [cit. 2015-03-06]. Dostupné z: <<http://www.asg.sav.sk/gfsb/v081/gfsb080109.pdf>>.

SVOBODOVÁ, J., TUČEK, P., ONDRÁČKOVÁ, J. *Evaluace digitálních modelů reliéfu metodami statistické analýzy*. Geodetický a kartografický obzor, 2009, roč. 55/97, č. 9, s. 227-231. ISSN 0016-7096.

VOJTEK, D. *Cizí slova a pojmy. Studijní opory pro cvičení* [online]. Ostrava: VŠB – Technická univerzita Ostrava, c2006, [cit. 2015-03-21]. Dostupné z: <<http://gis.vsb.cz/vojtek/index.php?page=dict/index>>.

VOŽENÍLEK, Vít a kol. *Integrace GPS/GIS v geomorfologickém výzkumu*. 1. vyd. Olomouc: Univerzita Palackého, 2001, 185 s. ISBN 80-244-0383-8.

VOŽENÍLEK, Vít. *Diplomové práce z geoinformatiky*. Olomouc: Vydavatelství Univerzity Palackého, 2002. 31 s. ISBN 80-244-0469-9.

WILSON, J. P., GALLANT, J. C.: *Terrain Analysis – Principles and applications*. John Wiley and Sons, New York, 2000, 433 s. ISBN 0-471-32188-5.

WOOD, J. D. *The geomorphological characterisation of digital elevation models* [online]. Leicester: Geography Department, University of Leicester, UK, c1996. [cit. 2015-02-03]. 456 s. Ph.D. thesis, Dostupné z: <http://www.geog.le.ac.uk/jwo/research/dem_char/thesis/>.

PŘÍLOHY

SEZNAM PŘÍLOH

Vázané přílohy:

- Příloha 1 Tabulky hodnot Spearmanova korelačního koeficientu a výsledné hodnoty OLS regrese – interpolační metoda: běžný kriging - semivariogram - sférický
- Příloha 2 Vizualizace prostorové analýzy reziduí OLS (matematický model)
- Příloha 3 Grafy závislosti reziduí OLS na parametrech (matematický model)
- Příloha 4 Vizualizace prostorové analýzy reziduí GWR (matematický model)
- Příloha 5 Tabulky OLS regrese při odstranění vysokého faktoru VIF a vzájemných korelací mezi parametry (laser scan)
- Příloha 6 Vizualizace prostorové analýzy reziduí OLS (laser scan)
- Příloha 7 Vizualizace prostorové analýzy reziduí GWR (laser scan)

Volné přílohy

- Příloha 8 DVD

Příloha 1 Tabulky hodnot Spearmanova korelačního koeficientu a výsledné hodnoty OLS regrese – interpolační metoda: běžný kriging - semivariogram – sférický

Tab. 1 Hodnoty Spearmanova korelačního koeficientu mezi hodnotami chyb a hodnotami základních zkoumaných parametrů – běžný kriging

Matematický terén 1 x 1 km	Zkoumané parametry									
		nadmořská výška	převýšení	vzdálenost	sklon	orientace	orientace reklasifikovaná	křivost celková	křivost planární	křivost profilová
<i>Vrstevnice</i>										
1 m	chyby	0,362	0,398	-0,005	0,008	0,004	0,034	-0,396	-0,373	0,398
3 m	chyby	0,473	-0,038	0,009	-0,034	-0,055	-0,048	-0,052	0,022	0,110
5 m	chyby	0,471	0,010	0,006	0,015	-0,006	-0,007	0,027	0,068	0,020
<i>Náhodně rozmístěné body</i>										
5%	chyby	0,205	0,040	-0,030	0,044	-0,018	-0,023	0,160	0,116	-0,130
10%	chyby	0,527	-0,009	0,000	-0,007	0,005	-0,002	0,475	0,430	-0,407
55%	chyby	-0,498	0,070	0,017	0,078	0,013	0,004	-0,535	-0,602	0,600
<i>Pravidelně rozmístěné body</i>										
15 m	chyby	0,513	-0,026	0,012	-0,028	-0,009	0,005	0,560	0,663	-0,489
25 m	chyby	0,485	-0,026	0,010	-0,027	-0,011	-0,010	0,510	0,487	-0,640
50 m	chyby	0,478	-0,044	0,007	-0,048	-0,039	-0,030	0,593	0,490	-0,459

Těsnost závislosti: nulová nízká mírná význačná

Tab. 2 Výsledné hodnoty prostého členu α , regresních koeficientů a koeficientu determinace Adjusted R^2 získané pomocí vícenásobné lineární OLS regrese - běžný kriging

Matematický terén 1 x 1 km	α	Regresní koeficienty nezávisle proměnných									Ad. R^2
		β_1 nadmořská výška	β_2 převýšení	β_3 vzdálenost	β_4 sklon	β_5 orientace	β_6 orientace reklasifikovaná	β_7 křivost celková	β_8 křivost planární	β_9 křivost profilová	
<i>Vrstevnice</i>											
1 m	0,008	0,008	0,001	-0,004	-0,001	0,000	0,002	0,009	-0,049	0,016	0,179
3 m	0,043	0,040	0,021	0,002	-0,008	0,000	-0,008	0,016	-0,023	0,088	0,177
5 m	-0,184	0,065	0,052	0,006	0,004	0,000	0,000	0,143	-0,154	0,245	0,162
<i>Náh. rozmís. body</i>											
5%	-0,041	0,011	0,194	0,001	-0,035	0,000	-0,005	0,347	-0,618	0,384	0,130
10%	-0,092	0,029	0,082	0,001	-0,007	0,000	0,004	0,509	-0,437	0,430	0,195
55%	-0,003	0,000	-0,019	0,000	0,005	0,000	-0,001	0,019	-0,074	0,069	0,285
<i>Prav. rozmís. body</i>											
15 m	-0,035	0,011	0,095	0,000	-0,019	0,000	0,003	0,415	0,041	0,185	0,309
25 m	-0,071	0,032	0,176	-0,001	-0,030	0,000	0,002	0,346	-0,039	0,519	0,268
50 m	-0,428	0,173	0,205	0,000	-0,171	0,000	0,010	0,322	-0,215	0,455	0,151

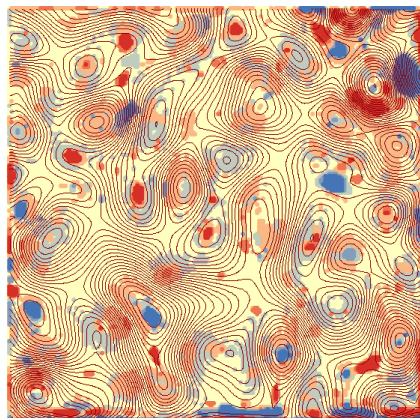
vysoký faktor VIF (Variance Inflation Factor) VIF > 7,5

Příloha 2 Vizualizace prostorové analýzy reziduí OLS (matematický model)

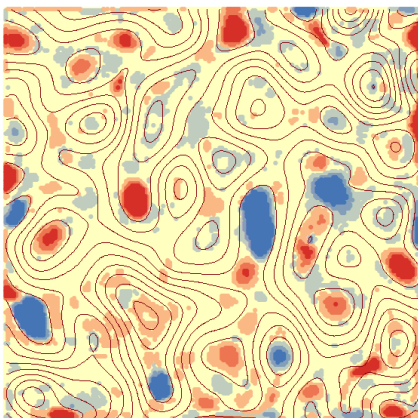


Vrstevnice:

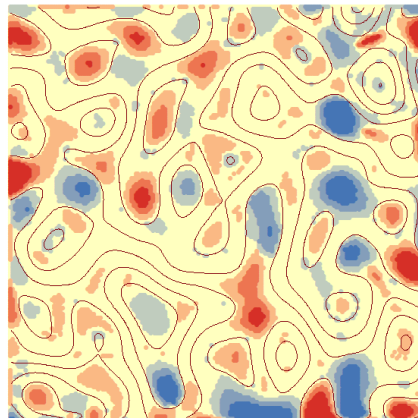
1 m



3 m

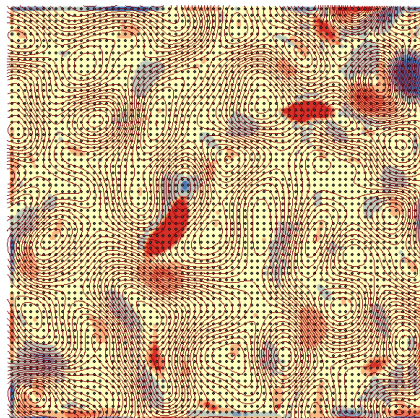


5 m

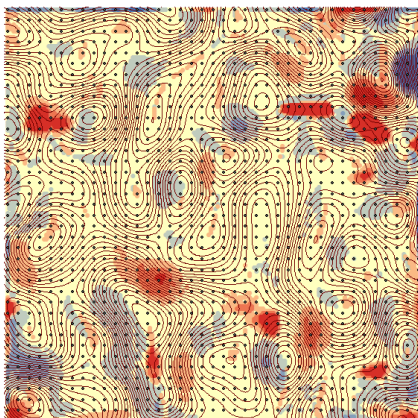


Pravidelně rozmístěné body:

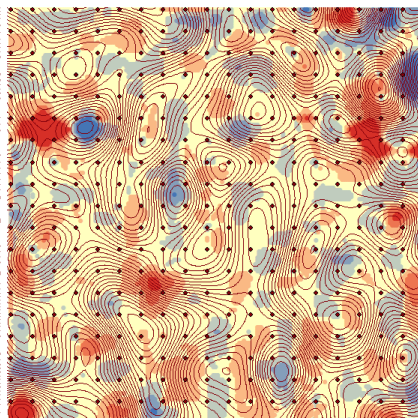
15 m



25 m

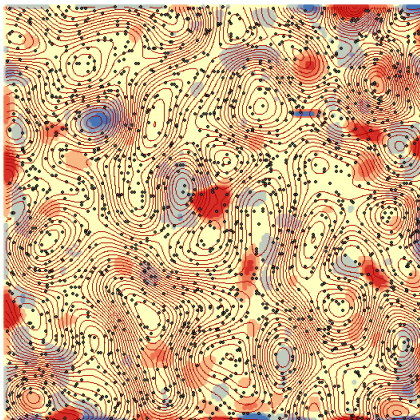


50 m

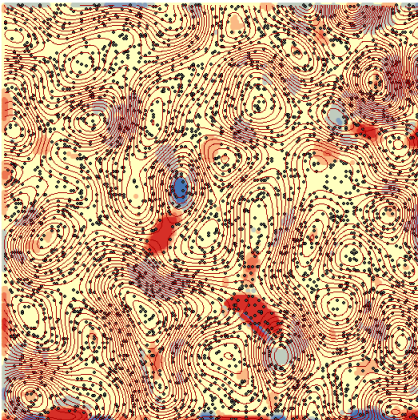


Náhodně rozmístěné body:

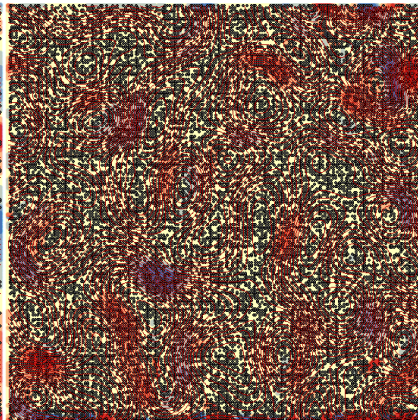
5 %



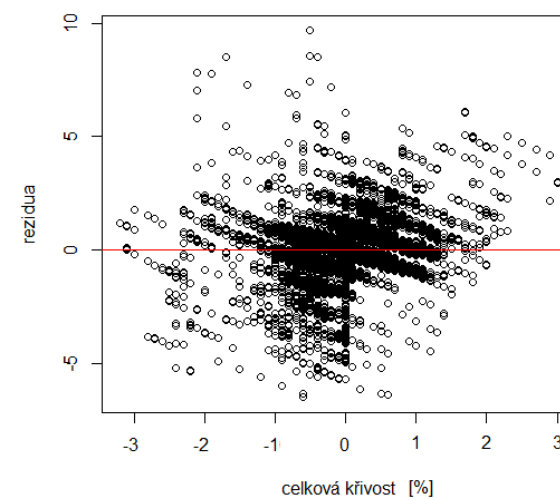
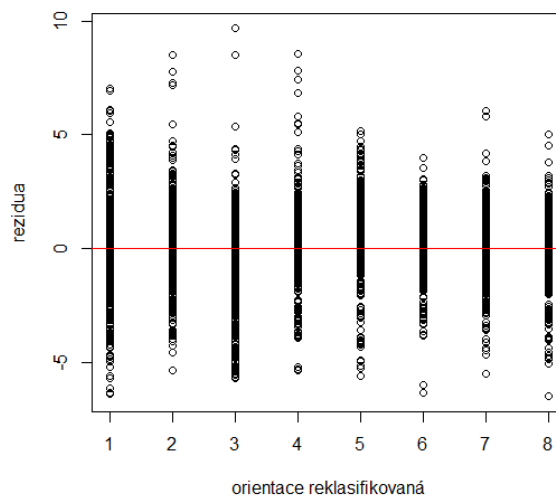
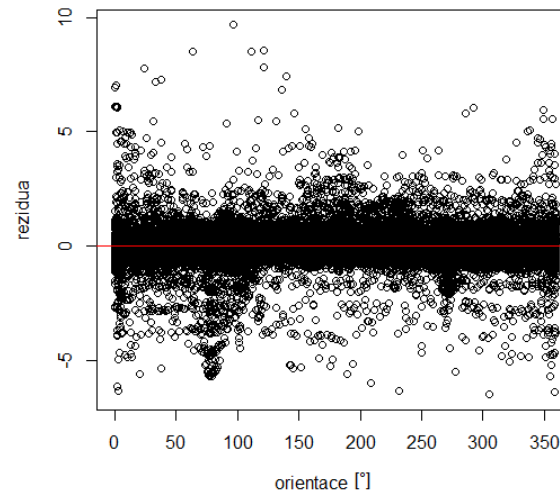
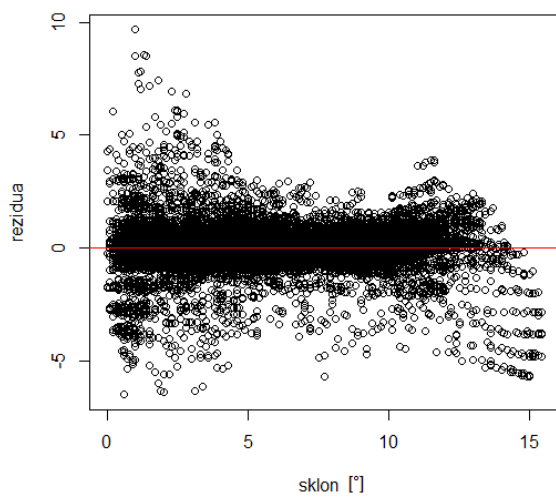
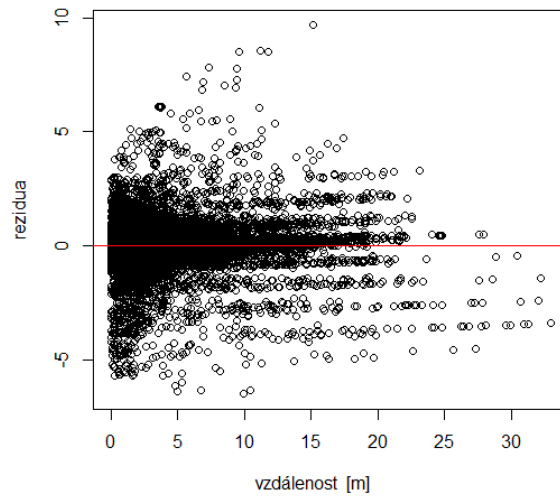
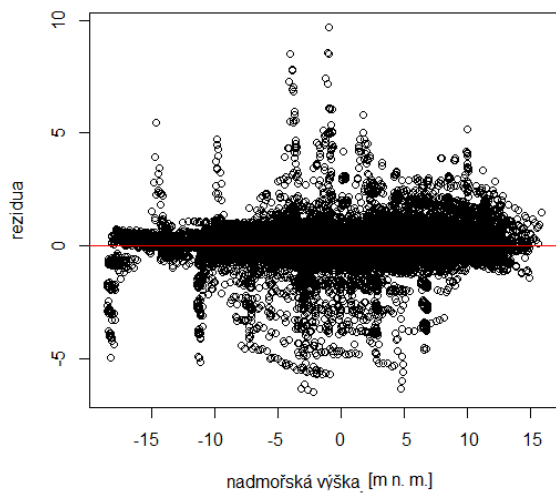
10 %



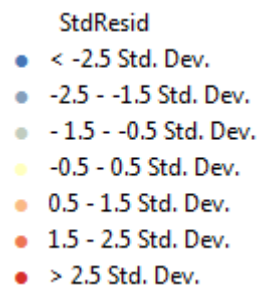
55 %



Příloha 3 Grafy závislosti reziduí OLS na parametrech (matematický model)



Příloha 4 Vizualizace prostorové analýzy reziduí GWR (matematický model)

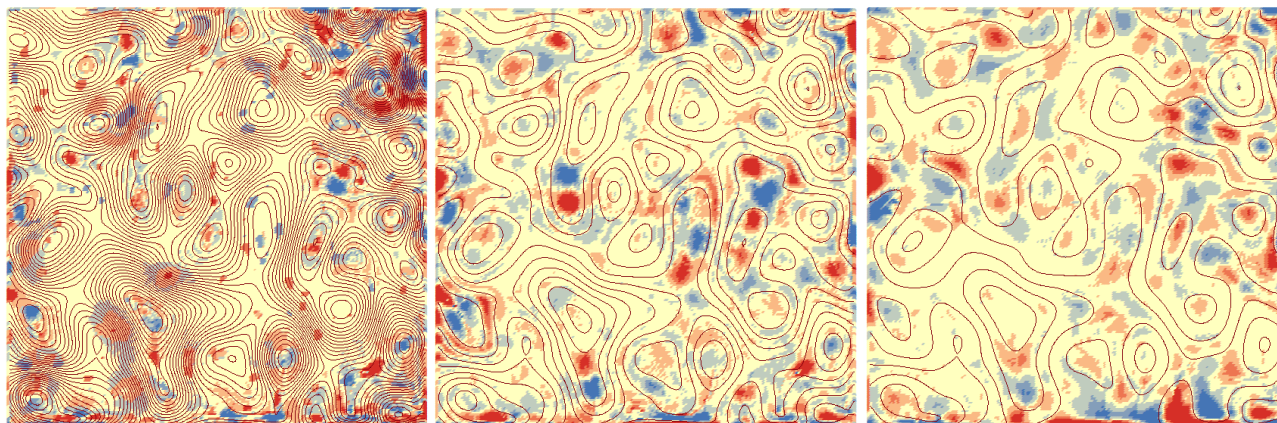


Vrstevnice:

1 m

3 m

5 m

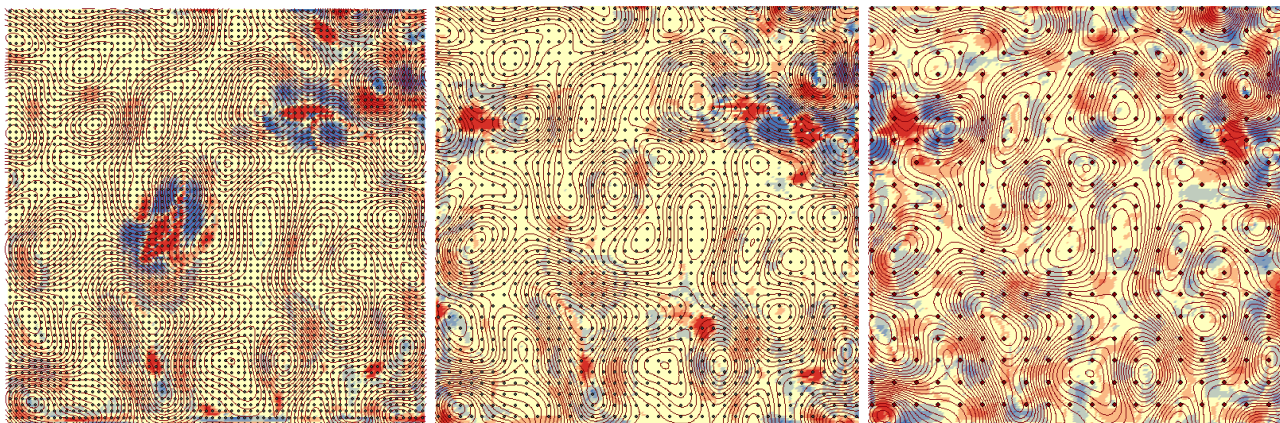


Pravidelně rozmístěné body:

15 m

25 m

50 m

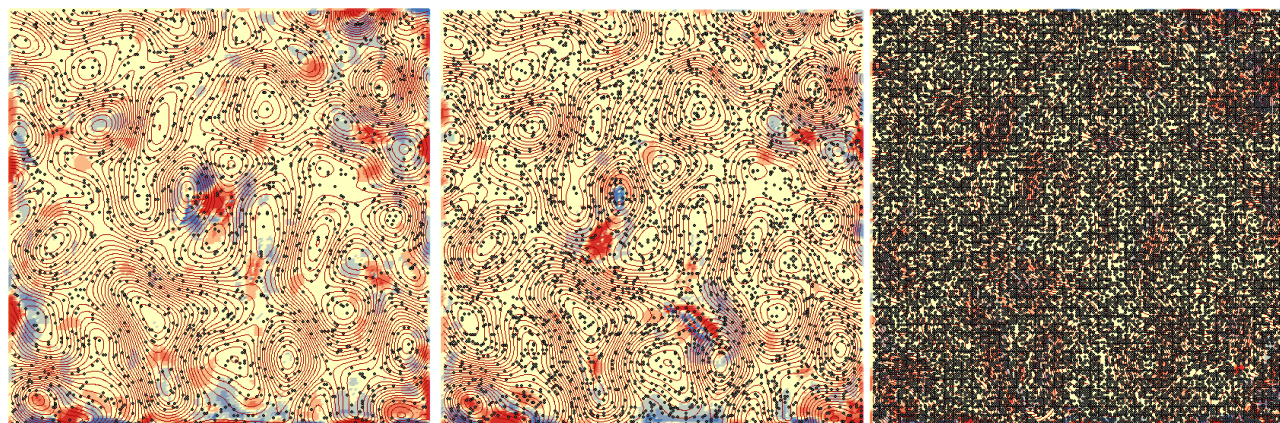


Náhodně rozmístěné body:

5 %

10 %

55 %



Příloha 5 Tabulky OLS regrese při odstranění vysokého faktoru VIF a vzájemných korelací mezi parametry (laser scan)

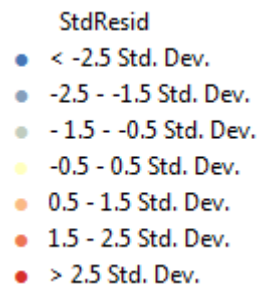
Tab. 1 Výsledné hodnoty prostého členu α , regresních koeficientů a koeficientu determinace Adjusted R^2 po odstranění parametrů na základě vysokého faktoru VIF získané pomocí vícenásobné lineární OLS regrese.

Laser-scan 1,1 x 1,4 km	α						Ad. R^2
		β_1 vzdálenost	β_2 orientace	β_3 orientace reklasifikovaná	β_4 sklon x profilová křivost	β_5 orientace x planární křivost	
<i>Vrstevnice</i>							
1 m	-0,041	0,000	0,000	0,006	-0,026	0,000	0,268
3 m	0,003	0,004	0,000	-0,006	-0,021	0,000	0,098
5 m	-0,049	0,010	0,000	-0,016	-0,014	0,001	0,066
<i>Náh. rozmís. body</i>							
5%	-0,240	-0,001	0,000	0,062	-0,028	0,001	0,060
10%	-0,104	0,002	0,000	0,020	-0,027	0,000	0,073
55%	-0,031	0,001	0,000	0,009	-0,022	0,000	0,193
<i>Prav. rozmís. body</i>							
15 m	-0,040	0,002	0,000	0,011	-0,022	0,000	0,083
25 m	-0,183	0,002	0,000	0,034	-0,028	0,000	0,068
50 m	-0,149	0,000	0,000	0,029	-0,017	0,002	0,020

Tab. 2 Výsledné hodnoty prostého členu α , regresních koeficientů a koeficientu determinace Adjusted R^2 po odstranění jednoho z dvojice parametrů, u kterých byla prokázána příliš silná korelace získané pomocí vícenásobné lineární OLS regrese.

Laser-scan 1,1 x 1,4 km	α					Ad. R^2
		β_1 vzdálenost	β_2 sklon	β_3 orientace reklasifikovaná	β_4 křivost celková	
<i>Vrstevnice</i>						
1 m	-0,044	0,001	0,000	0,010	0,163	0,203
3 m	-0,030	0,005	0,003	-0,001	0,130	0,068
5 m	-0,086	0,010	0,003	0,001	0,141	0,052
<i>Náh. rozmís. body</i>						
5%	-0,292	-0,001	0,006	0,062	0,228	0,057
10%	-0,149	0,002	0,004	0,027	0,154	0,059
55%	-0,042	0,001	0,001	0,009	0,117	0,123
<i>Prav. rozmís. body</i>						
15 m	-0,072	0,001	0,003	0,011	0,124	0,063
25 m	-0,224	0,002	0,004	0,041	0,141	0,053
50 m	-0,331	-0,001	0,018	0,048	0,245	0,024

Příloha 6 Vizualizace prostorové analýzy reziduí OLS (laser scan)

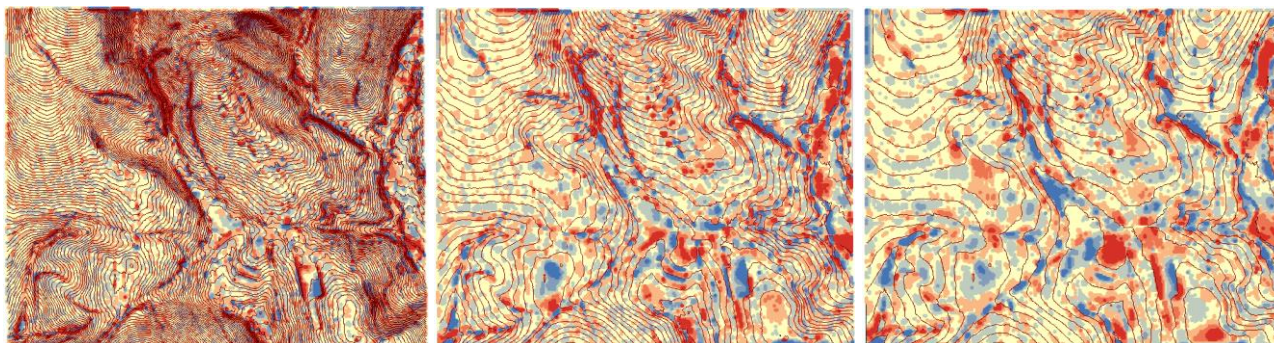


Vrstevnice:

1 m

3 m

5 m

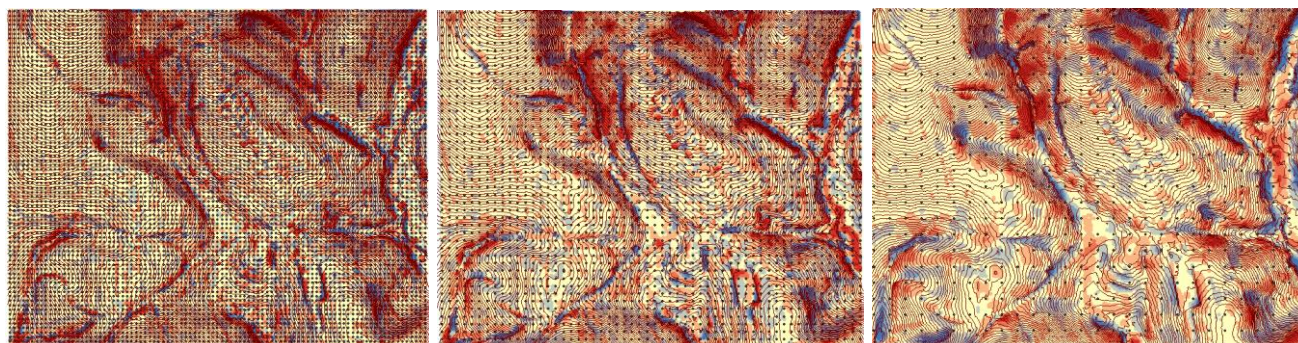


Pravidelně rozmístěné body:

15 m

25 m

50 m

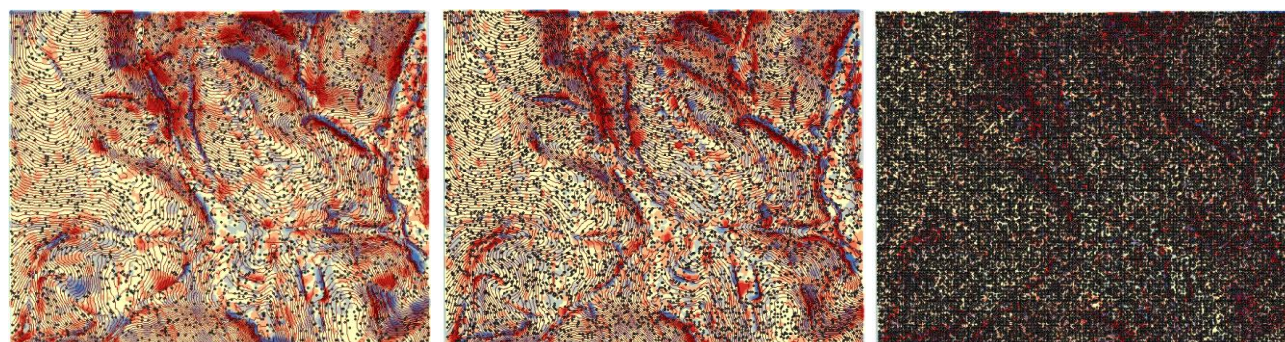


Náhodně rozmístěné body:

5 %

10 %

55 %



Příloha 7 Vizualizace prostorové analýzy reziduí GWR (laser scan)

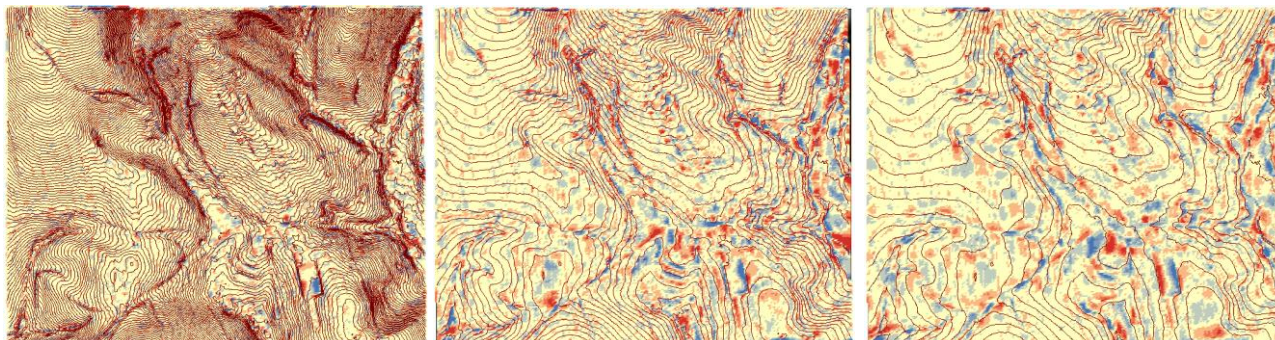


Vrstevnice:

1 m

3 m

5 m

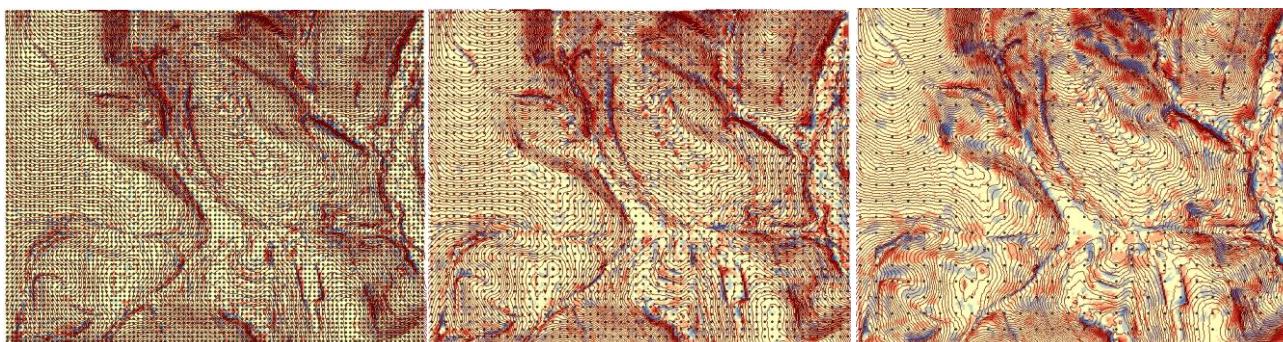


Pravidelně rozmístěné body:

15 m

25 m

50 m



Náhodně rozmístěné body:

5 %

10 %

55 %

